2019

Identifying an Optimal Bald Eagle Monitoring Program for Southwest Alaska National Parks

Rebecca Kolstrom
South Dakota State University

IDENTIFYING AN OPTIMAL BALD EAGLE MONITORING PROGRAM FOR SOUTHWEST ALASKA NATIONAL PARKS

BY

REBECCA KOLSTROM

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Biological Sciences

South Dakota State University

2019

IDENTIFYING AN OPTIMAL BALD EAGLE MONITORING PROGRAM FOR SOUTHWEST ALASKA NATIONAL PARKS

REBECCA KOLSTROM

This thesis is approved as a creditable and independent investigation by a candidate for the Master of Science degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Larry Gigliotti, Ph.D.
Thesis Advisor Date

Tammy Wilson, Ph.D.
Thesis Advisor Date

Michele Dudash, Ph.D.
Head, Department of Natural Resource Management Date

Dean, Graduate School Date

ACKNOWLEDGMENTS

I would first like to thank my advisors, Dr. Larry Gigliotti and Dr. Tammy Wilson for their guidance throughout my degree. They have been helpful and accessible throughout every step of this process. Both have been excellent mentors over the past several years and I have learned so much from each of them. I would also like to thank Dr. Corey Shelsta for being a part of my committee as a Graduate Faculty Representative. I am grateful to my funding sources and supporting institutions, the U.S. Geological Survey and South Dakota Cooperative Fish and Wildlife Research Unit, National Park Service, and South Dakota State University.

I would also like to thank each member of the Bald Eagle Expert Panel, without whom this project would not have been possible: A. Miller, M. MacCluskie, J. Putera, B. Mangipane, J. Schmidt, C. McIntyre, D. Payer, E. Dowd Stukel, L. Gigliotti, M. Hake, S. Kim, S. Lewis, T. Hamon, T. Shepherd, T. Wilson, T. Mullet, A. Matz, J. Reynolds, and L. Phillips. I greatly appreciate the time and commitment of these individuals and cannot thank them enough for providing the information necessary to complete this project.

I owe a huge thank you to my friends and family. To my parents and Timmy, thank you for always supporting me. A special thanks to the NRM graduate students, who have become close friends, for creating a great community at SDSU. Thank you to Austin for being incredibly supportive and patient. I am also grateful to the faculty and staff at SDSU, especially those in the main office, who are always willing to help. Finally, I would like to acknowledge all those who have helped me to succeed. From those at Penn State who encouraged my passions, to coworkers and crew leads who have helped me explore this field, to those who have encouraged me at SDSU – thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

IDENTIFYING AN OPTIMAL BALD EAGLE MONITORING PROGRAM FOR SOUTHWEST ALASKA NATIONAL PARKS

REBECCA KOLSTROM

2019

The Southwest Alaska Inventory and Monitoring Network includes bald eagle monitoring as part of their Vital Signs Monitoring Plan. Lake Clark National Park and Preserve, Katmai National Park and Preserve, Kenai Fjords National Park, and Wrangell – St. Elias National Park and Preserve monitor bald eagles annually, albeit slightly differently among parks. Since monitoring decisions involve multiple objectives and stakeholders, there was a need for a structured approach to identify an optimal monitoring program.

We used a structured decision making process and an iterative, four-round Delphi Process to collect information about long-term bald eagle monitoring from experts. We collected information about important stressors to bald eagles, and information about various monitoring metrics. We also held an in-person meeting with members of the expert panel to designate fundamental objectives for decisions about the long-term bald eagle monitoring, which are: 1) Minimize cost; 2) Minimize effort; 3) Maximize amount of accurate information collected about bald eagles; 4) Maximize the ability to detect change in bald eagle populations.

We used a consequence table to compare monitoring metrics and reduce the list of metrics to consider for the program. Panelists weighted the four fundamental objectives

by importance using a swing-weighting technique. Objectives weights are calculated using averages of panelist response: Maximize accurate information: 33.1%; Maximize ability to detect change: 32.3%; Minimize effort: 17.6%; Minimize cost: 17.1%.

A Bayesian Decision Net, which uses linear value modeling, compares alternative monitoring scenarios using information collected during the Delphi Process and the weight of fundamental objectives to determine the most optimal scenario. Our model identified a comprehensive monitoring scenario, which includes all feasible monitoring metrics, as the most optimal decision, followed by the current monitoring scenario. We performed a cross-stakeholder sensitivity analysis and an additional sensitivity analysis by varying objective weights. We also performed a sensitivity analysis using a two-function decision model, combining similarly weighted objectives into two objectives. We found that the cost and effort of the comprehensive monitoring scenario must be 4.4 times greater than the cost and effort of the current scenario, for the current monitoring scenario to become the most optimal decision.

CHAPTER 1

INTRODUCTION TO BALD EAGLE MONITORING IN SOUTHWEST ALASKA NATIONAL PARKS

BRIEF REGULATORY HISTORY OF BALD EAGLES

Bald eagles are currently protected by various laws and acts (U.S. Fish and Wildlife Service 2015). Populations of bald eagles notoriously decreased during the mid- to late-20th century due to the insecticide DDT. Surveys performed by the National Audubon Society found poor nesting success and few nesting pairs of bald eagles, which validated anecdotal evidence of population declines (Carson 1962, U.S. Fish and Wildlife Service 2009). DDT was banned in 1972 and bald eagles were listed as endangered under the Endangered Species Act in 43 of the 48 contiguous states. In 1995, following recovery of populations, bald eagles were reclassified as threatened throughout the lower 48 and were officially delisted in 2007 (U.S. Fish and Wildlife Service 2009). Bald eagles are monitored following protocols outlined in the Bald Eagle Post-Delisting Monitoring Plan, and are protected under the Bald and Golden Eagle Protection Act (1940), the Migratory Bird Treaty Act (1918), and the Lacey Act (1900) (U.S. Fish and Wildlife Service 2009;2015).

BALD EAGLE BIOLOGY

Bald eagles span most of North America, inhabiting areas near aquatic resources with suitable nesting locations – most commonly, in the tops of large trees; ideal habitats contain abundant food sources, are subject to little disturbance by humans, and are often comprised of old growth forests (Stalmaster 1987, Buehler 2000, Suring 2008). Large breeding populations of bald eagles in the United States are located in Alaska, the Great Lakes states, Florida, the Pacific Northwest, Maine, and the area surrounding the Chesapeake Bay (Stalmaster 1987). Breeding bald eagles have not been recorded outside of North America and, although some populations migrate, most are thought to remain on or in the vicinity of their breeding territory year-round (Stalmaster 1987, Buehler 2000). A map of bald eagle range is shown in Figure 1.1, from Buehler (2000).

High quality habitat is important to bald eagles, as they are known to consistently favor landscape features conducive to nesting, perching, roosting, and foraging (Suring 2008). Bald eagles prefer open spaces near open water to exploit their prey sources, especially since they are unable to fly through dense forest stands to kill and eat prey. Habitat selection is also based on consistently available prey (Stalmaster 1987). Bald eagles can be territorial, and territorial behaviors, such as threatening vocalizations, chasing, and perching in visible areas, is more common during the breeding season (Stalmaster 1987, Buehler 2000). Nesting sites are chosen within a close proximity to water and are almost always located in trees (frequently the tallest trees in the forest stand) (Stalmaster 1987). However, ground nests may be located on sea stacks, cliffs, or other prominent landscape features where forested areas are more scarce, such as in western Alaska and the Aleutian Islands (Suring 2008).

Various raptor species frequently reuse nests constructed in previous years. A study of forest raptors documented higher incidences of establishment in old territories than selection of new territorial settlements (Jiménez-Franco et al. 2014). Bald eagle pairs, which are generally stable and assumed to mate for life, exhibit high degrees of fidelity to nest sites (Stalmaster 1987, Jenkins and Jackman 1993). Since habitat is an important variable when selecting nest sites, with most nest sites being chosen near water resources for foraging opportunities, high rates of nest reuse may indicate that these old nests may be valuable resources as well as important territorial cues (Hansen and Hodges 1985, Suring 2008, Jiménez-Franco et al. 2014). Bald eagles may have more than one nest in their territory, and in populations of Alaskan bald eagles, successful nests have a higher observed probability of being reused in the following year, while unsuccessful nests are less likely to be used for nesting the following year (Gende et al. 1997, Bailey et al. 2008, Wilson et al. 2018). Nest reuse may provide reproductive benefits. A study of forest raptors reported higher probability of breeding success or more fledglings in pairs of birds reusing nests (Jiménez-Franco et al. 2014).

Bald eagles are opportunistic feeders that will eat a wide variety of prey, though appear to select fish over other food sources, which include birds and mammals (Stalmaster 1987). Salmon are particularly important for bald eagles and eagle foraging habitat is often designated by areas that can support populations of these fish (Stalmaster 1987, Buehler 2000). Bald eagles generally forage through three methods. These are stealing prey from other animals, scavenging, and hunting. While eagles forage using all of these methods, studies have shown that bald eagles appear to favor stealing to scavenging and hunting, and favor scavenging to hunting (Stalmaster 1987). Eagles often lose possession of prey to other eagles, and it appears that they can compare their hunger level and body condition to that of other eagles to determine whether stealing will be successful (Hansen et al. 2008). Eagles may steal from conspecifics (pirating) or from other species (kleptoparasitism). When hunting for prey, eagles may hunt while in flight, from a perch, wading, from the ground, or cooperatively with other eagles (Stalmaster 1987).

Bald eagles are a long-lived species, with a record longevity of 28 years recorded in the wild (Buehler 2000). Survivorship is largely dependent upon food availability, and bald eagles are scarce in areas with limited food availability (Hansen et al. 2008). Hansen et al. (2008) states that "food is the web that interconnects virtually every aspect of bald eagle ecology." The juvenile life history stage corresponds to the lowest survival of eagles. If bald eagles can survive the juvenile life stage, survivorship is generally quite high (Stalmaster 1987).

Breeding behavior begins in the spring with vocal displays, chase displays, roller-coaster flight, and cartwheel display (Stalmaster 1987, Gende 2008). The cartwheel display is a dramatic flight, where eagles fly to high altitudes and lock talons, tumbling toward the ground and breaking off just before collision (Buehler 2000). In the pre-laying stage of breeding, eagles begin to spend larger amounts of time in territories and begin to build or rebuild nests, with nest building usually commencing one to three months before laying eggs (Buehler 2000, Gende 2008). Bald eagles are considered immature until they are about four years old, when they display their definitive, characteristic plumage and are able to begin breeding. Most commonly, they lay two eggs per clutch and produce only one brood per season, although larger clutches are occasionally observed (Buehler 2000). Eggs are generally laid over the span of several days, with incubation beginning immediately and (at first) performed nearly 24 hours a day (Gende 2008). Once eggs are laid, both males and females incubate the eggs, alternating throughout the day and sharing nesting duties more evenly than other birds of prey (Cain 2008). Both sexes have brood patches, however the female brood patch is generally larger and more developed than the male brood patch (Buehler 2000). Eagles perform various behaviors during incubation, such as egg turning and nest raking (Cain 2008). Both male and females deliver prey items to young eagles in the nest. In three studied nests in Alaska, fledging occurred between 80 and 86 days post-hatching (Cain 2008). Generally, the time span from the first egg being laid until the last young fledging from the nest ranges from 16 to 18 weeks (Stalmaster 1987), and young bald eagles are known to leave the nest before the completion of flight feather growth (Bortolotti 1984).

Productivity and reproductive success are dependent on factors which include food availability, weather conditions, contaminants, and human disturbances. Food availability, as influenced by the quality of habitat, can affect many reproductive phases in eagles (Hansen and Hodges 1985, Hansen 1987, Hansen et al. 2008). Abundant food sources that are readily available to eagles in prelaying and incubation periods have been observed to potentially affect nest productivity measures (Gende et al. 1997). Early laying dates and enhanced offspring survival have also been observed in the presence of high food availability and supplemental feedings (Hansen 1987). Additionally, weather variables were shown to have an impact on almost all activities relating to nesting in Alaskan bald eagles. Harsh weather may influence the time budgets of nesting eagles, which can increase incubation demands on nesting pairs and reduce time allotted to foraging (Cain 2008). Prey availability and weather conditions can also interact to influence hatching dates, reproductive rates of eagles, and to limit the productivity of raptor populations (Steenhof et al. 1997).

Various stressors have affected bald eagles throughout their range. In the contiguous United States, specifically in the Great Lakes region, availability of habitat, human disturbance, and environmental contaminants are considered the largest influences on bald eagle productivity (Bowerman et al. 1995). Bald eagles in the lower 48 states have historically been subjected to acute effects from environmental contaminants and direct human disturbance, which also affect bald eagle productivity. Productivity of eagles was diminished significantly in the presence of high concentrations of DDE, total PCBs, and other organochlorine pesticides (Dykstra et al. 2001). More specifically, DDE was identified to have a negative relationship with reproduction levels of bald eagles at a population level. DDT was banned by the Environmental Protection Agency in the United States in 1972 (U.S. Fish and Wildlife Service 2009) and subsequent studies showed significantly increased reproduction (Grier 1982).

Direct disturbance, defined as any human activity that causes deviations in bald eagle behavior patterns, can affect normal activity patterns, spatial use patterns and activity budgets of bald eagles. This includes changes in behaviors (type, duration, and frequency) and changes in the responses of eagles to environmental factors, which can negatively affect reproduction (Steidl and Anthony 2000, Cain 2008, Fraser and Anthony 2008). Both direct and indirect disturbances affect bald eagles, and exposure to these disturbances during the breeding season has the potential to drive eagles to find new nest sites (Fraser and Anthony 2008). Especially in the contiguous states, increased development for human use has impacted eagle habitat, including suitable nesting trees (US Army Corps of Engineers 2008). Bald eagles, like many other plants and animals, are affected and limited by climate change and various climatic factors (Bennett et al. 2006).

BALD EAGLES IN ALASKA

Bald eagles in Alaska have a unique and storied history. Populations experienced a decline from 1917 to 1952, when a bounty was placed on the bald eagle. At this time, the birds were not protected by the Bald Eagle Protection Act, and only received protection when Alaska officially became a state in 1959 (Cegelske 2008, Jacobsen 2008b). A 1920 edition of The Valdez Miner newspaper claims that "The eagle is a curse to the rest of the animal kingdom and the sooner it is exterminated, the better off the game will be." Even the Deputy U.S. Commissioner of Fisheries declared the bald eagle as a destructive threat to the salmon population (DeArmond 2008). During the period of legalized persecution, at least 128,273 bald eagles were cashed in for bounty, according to data from the Alaska Territorial Treasurer (Robards and King 1966) and 50 cents were provided to hunters who presented the feet of an eagle (DeArmond 2008). Although the bounty was lifted after a bill was passed by lawmakers, killing of eagles by electrocution, oil spills, trapping, and shooting remained a problem (Robards and King 1966, Cegelske 2008, DeArmond 2008, Jacobsen 2008b). However, following the removal of the bounty, attitudes toward bald eagles began to change and fish remained abundant, allowing populations to recover (Robards and King 1966, Jacobsen 2008b).

Following the *Exxon Valdez* oil spill in 1989, efforts were made to determine the effects of the oil spill on bald eagle populations. There is variation in bald eagle productivity and nest success data from Alaska prior to the oil spill, making it difficult to determine the extent of the impact. However, studies found that the only area that exhibited a large reproductive failure event was western Prince William Sound. Overall, survival of bald eagles remained high, as long as hatching occurred (Bowman et al. 1997, Bernatowicz et al. 2008). Alaska Department of Fish and Game currently reports the population of bald eagles in the state around 30,000 birds (Alaska Department of Fish and Game 2017). The state of Alaska also hosts the world's largest congregation of bald eagles each fall in the Chilkat River drainage (Jacobsen 2008b). Bald eagles remained so abundant in Alaska that young bald eagles were translocated from Alaska to states in the lower 48 and successfully released into the wild to supplement natural populations and recovery efforts (Jacobsen 2008a).

STUDY AREA

The Southwest Alaska Inventory and Monitoring Network (SWAN) is a collection of National Parks and Monuments located on the shelf of the North American Plate in an extremely geologically active region. SWAN consists of Katmai National Park and Preserve (KATM), Kenai Fjords National Park (KEFJ), Lake Clark National Park and Preserve (LACL), Alagnak Wild River (ALAG), and Aniakchak National Monument and Preserve (ANIA) (Bennett et al. 2006). SWAN is "dedicated to providing the scientific foundation for effective, long-term protection and management of natural resources in five units of the national park system" (National Park Service 2018). Together, the five units of SWAN parks cover 9.4 million acres, include three Alaska climatic zones and 11 ecoregions, and contain nearly one third of the National Park System's marine coastline. This coastal habitat comprises 1,200 miles of diverse coastline in the northern Gulf of Alaska (Bennett et al. 2006, National Park Service 2018). This provides extensive habitat for nesting bald eagles. Specifically, LACL, KATM, and KEFJ are home to populations of breeding bald eagles, which are monitored annually (Wilson et al. 2017). Wrangell – St. Elias National Park and Preserve is part of the Central Alaska Inventory and Monitoring Network, and is also home to breeding bald eagles. The parks included in this study area are displayed in Figure 1.2.

BALD EAGLE MONITORING IN SOUTHWEST ALASKA NETWORK

Bald eagles are currently monitored by the Southwest Alaska Inventory and Monitoring Network as part of their Vital Signs Monitoring Plan (Bennett et al. 2006). The plan specifies the definition of vital signs monitoring as "the collection and analysis of repeated observations or measurements to evaluate ecological changes in the condition of park resources." Long-term monitoring is an important component of the National Park Service mission, as it aids in the understanding of natural resources and ecosystem dynamics (Bennett et al. 2006). The goals of the National Park Service Vital Signs Monitoring are as follows:

1. Determine status and trends in selected indicators of the condition of park ecosystems to allow managers to make better-informed decisions and to work more effectively with other agencies and individuals for the benefit of park resources.

2. Provide early warning of abnormal conditions of selected resources to help develop effective mitigation measures and reduce costs of management.

3. Provide data to better understand the dynamic nature and condition of park ecosystems and to provide reference points for comparisons with other, altered environments.

4. Provide data to meet certain legal and congressional mandates related to natural resource protection and visitor enjoyment.

5. Provide a means of measuring progress toward performance goals.

Bald eagles were selected as a terrestrial animal signal of biological integrity, because of their role as a keystone predator on avian and fish populations in the system. Their goal was stated to: "Estimate long-term trends in nest occupancy and productivity from a random sample of bald eagles nesting along interior rivers/lakes and marine coastlines of SWAN parks." (Bennett et al. 2006).

Monitoring for bald eagles in the Southwest Alaska Network currently follows the protocol and standard operating procedures, published by the National Park Service in 2016 (Wilson et al. 2016). Bald eagles are affected by a variety of stressors, and can indicate changes in both the long- and short-term (Thompson et al. 2009). Lake Clark National Park and Preserve, Katmai National Park and Preserve, and Kenai Fjords National Park are currently addressing the following objectives with their monitoring program:

- "Estimate long-term trends in the abundance of bald eagle nests located within the sampled areas. Abundance is not directly observable, and will be modeled using data obtained using an estimator that uses observations from two observers."
- "Estimate long-term trends in the annual proportion of nests in which eagles attempt to reproduce (nest initiation). Nest initiation is not directly observable

and will be modeled using data obtained during two nest initiation surveys conducted in May."

- "Estimate long-term trends in annual nest productivity. Productivity is defined as the mean number of chicks produced per initiated nest, and is conditional on nest occupancy" (Wilson et al. 2016).

Wrangell – St. Elias National Park and Preserves also monitors bald eagles to determine nesting territory occupancy, nesting success, and population productivity (Putera and Miller 2018). Currently, the parks use slightly different methods to monitor for bald eagles. Parks use variations on both the double-observer and dual-frame methods. The double-observer method helps to account for imperfect detection when performing count surveys; in this case, for imperfect detection of nests. By using a survey technique that incorporates both a "primary" and "secondary" observer, detection probabilities can be estimated and applied to aerial nest counts. Double-observer methods are preferable to single-observer counts and counts where detection probability is not considered (Nichols et al. 2000). The dual-frame method refers to using both a list frame and an area frame (alternatively, the overlap domain and nonoverlap domain, respectively) to survey for eagle nests; the list frame consists of known eagle nests and is surveyed each year to determine activity at those known nests. The area frame is used to survey random transects or areas of potential eagle habitat for new nests (Haines and Pollock 1998, U.S. Fish and Wildlife Service 2009). Since list frames may be incomplete and area frames provide the opportunity for full coverage of the population in question, using both frames together can help to counterbalance the strengths and weaknesses of each (Haines and Pollock 1998). To generate population estimates, Alpizar-Jara et al. (2005) suggest the use of a screening estimator under the dual-frame method.

Kenai Fjords National Park has been conducting aerial surveys along the coast since 2009 using a double observer and dual frame method. Katmai National Park and Preserve has been surveying for bald eagles in the Naknek watershed since 2010, also using the double observer and dual frame method. Lake Clark National Park and Preserve, however, uses only a single observer and the list frame to conduct their surveys, but they have collected data from a majority of the potential bald eagle habitat in the park since 1992 (Wilson et al. 2016). Wrangell – St. Elias National Park and Preserve closely adheres to the SWAN Protocol for Lake Clark National Park and Preserve. Bald eagles have been monitored in WRST since 1987. WRST uses list frame monitoring to estimate territory-level metrics along the Copper River (Putera and Miller 2018). Discrepancies between sampling methods and definitions of monitoring metrics make it more difficult for park managers to make comparisons between bald eagle data and use bald eagles as indicators of environmental quality through overall population trends. Table 1.1, adapted from the protocol narrative, “Monitoring Bald Eagles in Southwest Alaska National Parks,” gives a summary of the sampling methods used for bald eagle surveys for each park in the Southwest Alaska Inventory and Monitoring Network and includes information about WRST.

A summary of the monitoring efforts by each park, which can be found in “Monitoring Bald Eagles in Southwest Alaska National Parks,” includes the area sampled and spatial design of the monitoring efforts (Wilson et al. 2017). KATM, KEFJ, and LACL all fly two surveys to investigate nest initiation and one survey to investigate productivity. However, for nest initiation surveys, KATM samples only the Naknek drainage (including the full shoreline and rivers around major lakes), KEFJ surveys all park coastline (using 23 coastal transects), and LACL surveys all known eagle habitat. WRST monitors bald eagles along the Copper River, Chitina River, Bremner River, Tanada Creek, and Copper and Tanada Lakes (Putera and Miller 2018).

Inference about the information collected is impacted in important ways, based on the observation methods used by the parks. For example, in the absence of the double observer method, measurements of absolute number of nests, used nests, and chick productivity will be biased low. For non-random observations of sample units, inference would be limited only to the units sampled, rather than the whole population; in the situation without the double observer method, previously described, measures must represent proportion of sampled nests used, proportion of sampled nests producing chicks, or the mean number of chicks per sampled nest.

PROJECT NEED

Management decisions, in a variety of contexts, rely heavily on the objectives that are formed by human values, budgetary concerns, and many other factors. However, if those objectives are not formed early in a monitoring program, time and other valuable resources may be wasted, conflict may arise, and management and conservation decisions may be less effective. Thus, the need for structured decision making in natural resource management is made very clear. The basis of structured decision making is formed around the principles of establishing specific and quantifiable monitoring objectives and various management alternatives, which lead to more focused and achievable management targets as well as transparency in the decision-making process (Conroy and

Peterson 2013). The roadmap for designing a biological monitoring program is very clear, and if followed correctly, can be an extremely effective tool in the decision making process (Reynolds et al. 2016). Figure 1.3 is adapted from the full roadmap presented in Reynolds et al. (2016) and outlines the steps to implement a successful monitoring program. Each successive step depends on the completion of the step before it, and monitoring programs should be re-evaluated often to ensure that they are achieving management goals. In the case of the Southwest Alaska Network of National Parks, several key steps were missed when creating a monitoring program for bald eagles. In the process of creating a monitoring plan, decision makers skipped over defining specific, measurable objectives and directly from sketching a conceptual model of the system to designing surveys and collecting data. Key portions of framing the problem and designing the objective were missed, and the parks now find themselves collecting data slightly differently, and unable to use their data as effectively and comprehensively as they would like. Although there is an abundance of data on bald eagle populations in the parks (Kenai Fjords: 2009-present; Katmai: 2010-present; Lake Clark: 1992-present), the data cannot be used broadly to make comparisons between parks and identify total population trends. Park scientists have developed a thorough understanding of assumptions, effective monitoring methods, and interpreting data, but it is now imperative that objectives are optimized through a structured decision process to maximize the applicability of data as well as minimize costs in the face of an uncertain budget.

This project seeks to fill gaps in the decision-making process, such as identifying alternatives, consequences, and tradeoffs, and the results of this project will be used to improve the long-term bald eagle monitoring program in the Southwest Alaska Network Inventory and Monitoring Program. Ultimately, the goals for this process are to combine stakeholder and expert opinions to form core management objectives for bald eagle populations in Southwest Alaska National Parks, and to standardize the long-term bald eagle monitoring protocol, so that data can be more broadly applied and effectively used.

The overarching objective that defines this project, is to uphold the mission of the National Parks Service by preserving and protecting bald eagle populations in SWAN National Parks. To achieve this broad goal, parks must ensure that bald eagles are maintaining reproduction rates and total population status – these are the means objectives that must be achieved to achieve the fundamental objective.

OBJECTIVES

Our research aims to use a formal structured decision making process to ensure that the bald eagle monitoring conducted by the parks is standardized and meets programmatic goals and objectives. This will be achieved using several smaller-scale objectives:

- Implement a Delphi process using online questionnaires to gather information and opinions from stakeholders.

  Stakeholders include National Park Service scientists and managers, eagle experts, and other interested parties. Through these surveys, we will gather information about important stressors for bald eagles (e.g. weather, disturbance, contaminants, etc.), how the population may change in response *to the stressors, and how those stressors and changes relate to population metrics.*

- Create a conceptual framework for bald eagles in Southwest Alaska National Parks.

  Stressors will be ranked in order of importance, and the most important metrics will be identified that measure population-level responses. We will map how those responses are linked to each stressor, and tie important metrics to fundamental decision objectives.

- Form a set of standardized set of fundamental objectives for the decision about the bald eagle monitoring program.

  These objectives will be formed by the Bald Eagle Expert Panel and will be based on the results of the questionnaires.

- Monitoring scenarios will be designed and analyzed, based on alignment with agreed-upon objectives.

  Monitoring scenarios will be designed to maximize information content of the metrics and minimize the resources required.

These data will be used to inform a decision about a standardized bald eagle management plan for SWAN that optimizes the fundamental objectives for the decision.

This process will identify core objectives, values, and expectations of bald eagle monitoring in Southwest Alaska National Parks. By involving National Park Service scientists and resource managers, and various other stakeholders, a reasonable consensus will be reached that will allow integrated sampling designs. This will ensure the sustainability of the bald eagle monitoring program in the Southwest Alaska Inventory and Monitoring Network. Thus, bald eagles can be monitored as a “vital sign,” as identified by the network, in an efficient manner that will optimize objectives and minimize costs.

LITERATURE CITED

Alaska Department of Fish and Game. 2017. Bald Eagle (Haliaeetus leucocephalus) Species Profile. <http://www.adfg.alaska.gov/index.cfm?adfg=baldeagle.main> Accessed 2018.

Alpizar-Jara, R., K. H. Pollock, and D. E. Haines. 2005. Mark-recapture estimators for dual frame population size of prominent nesting structures: the effect of uncertain detection probability. Environmental and Ecological Statistics 12:155-168.

Bailey, T. N., E. E. Bangs, W. W. Larned, A. J. Loranger, M. F. Portner, T. V. Schumacher, and E. A. Jozwiak. 2008. Nesting and wintering Bald Eagle population parameters on and adjacent to the Kenai National Wildlife Refuge, Alaska. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Bennett, A. J., W. L. Thompson, and D. C. Mortenson. 2006. Vital Signs Monitoring Plan, Southwest Alaska Network. National Park Service, Anchorage, AK.

Bernatowicz, J. A., P. F. Schempf, and T. D. Bowman. 2008. Bald Eagle productivity in Southcentral Alaska in 1989 and 1990 after the *Exxon Valdez* oil spill. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Bortolotti, G. R. 1984. Physical Development of Nestling Bald Eagles with Emphasis on the Timing of Growth Events. The Wilson Bulletin 96:524-542.

Bowerman, W. W., J. P. Giesy, D. A. Best, and V. J. Kramer. 1995. A Review of Factors Affecting Productivity of Bald Eagles in the Great Lakes Region: Implications for Recovery. Environmental Health Perspectives 103:51-59.

Bowman, T. D., P. F. Schempf, and J. I. Hodges. 1997. Bald Eagle Population in Prince William Sound after the Exxon valdez Oil Spill. The Journal of Wildlife Management 61:962-967.

Buehler, D. A. 2000. Bald Eagle (*Haliaeetus leucocephalus*), version 2.0. *in* P. G. Rodewald, editor. The Birds of North America. Cornell Lab of Ornithology, Ithaca, New York, USA.

Cain, S. L. 2008. Time Budgets and Behavior of Nesting Bald Eagles. Pages 73-94 *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Carson, R. 1962. Silent Spring. Houghton Mifflin, Cambridge, MA.

Cegelske, J. A. 2008. Law enforcement and the Bald Eagle Protection Act. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Conroy, M. J., and J. T. Peterson. 2013. Decision making in natural resource management: a structured, adaptive approach. Oxford: Wiley-Blackwell.

DeArmond, R. N. 2008. Shoot the damned things! Alaska's war against the American Bald Eagle. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Dykstra, C. R., M. W. Meyer, K. L. Stromborg, D. K. Warnke, W. W. Bowerman, and D. A. Best. 2001. Association of Low Reproductive Rates and High Contaminant Levels in Bald Eagles on Green Bay, Lake Michigan. Journal of Great Lakes Research 27:239-251.

Fraser, J. D., and R. G. Anthony. 2008. Human disturbance and Bald Eagles. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Gende, S. M. 2008. Perspectives on the breeding biology of Bald Eagles in Southeast Alaska. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Gende, S. M., M. F. Wilson, and M. Jacobsen. 1997. Reproductive success of bald eagles (*Haliaeetus leucocephalus*) and its association with habitat or landscape features and weather in southeast Alaska. Canadian Journal of Zoology 75:1595-1604.

Grier, J. W. 1982. Ban of DDT and Subsequent Recovery of Reproduction in Bald Eagles. Science 218:1232-1235.

Haines, D. E., and K. H. Pollock. 1998. Estimating the number of active and successful bald eagle nests an application of the dual frame method. Environmental and Ecological Statistics 5:245-256.

Hansen, A., E. L. Boeker, and J. I. Hodges. 2008. The population ecology of Bald Eagles along the Pacific Northwest coast. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Hansen, A. J. 1987. Regulation of Bald Eagle Reproductive Rates in Southeast Alaska. Ecology 68:1387-1392.

Hansen, A. J., and J. I. J. Hodges. 1985. High Rates of Nonbreeding Adult Bald Eagles in Southeastern Alaska. The Journal of Wildlife Management 49:454-458.

Jacobsen, M. J. 2008a. Removal of Alaskan Bald Eagles for translocation to states. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

_____. 2008b. The Status of the Bald Eagle in Southeast Alaska. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Jenkins, J. M., and R. E. Jackman. 1993. Mate and Nest Site Fidelity in a Resident Population of Bald Eagles. The Condor 95:1053-1056.

Jiménez-Franco, M. V., J. E. Martínez, and J. F. Calvo. 2014. Patterns of nest reuse in forest raptors and their effects on reproductive output. Journal of Zoology 292:64-70.

National Park Service. 2018. Southwest Alaska Parks and Partners. <https://www.nps.gov/im/swan/parks.htm>. Accessed 2018.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A Double-Observer Approach for Estimating Detection Probability and Abundance from Point Counts. The Auk 117:393-408.

Putera, J., and S. D. Miller. 2018. Protocol Implementation Plan for Bald Eagle Monitoring in Wrangell - St. Elias National Park & Preserve, Central Alaska Network. Report Natural Resource Report NPS/CAKN/NRR—2018/1752.

Reynolds, J. H., M. G. Knutson, K. B. Newman, E. D. Silverman, and W. L. Thompson. 2016. A road map for designing and implementing a biological monitoring program. Environmental Monitoring and Assessment 188:399.

Robards, F. C., and J. G. King. 1966. Nesting and Productivity of Bald Eagles: Southeast Alaska, 1966. Juneau, Alaska, USA.

Stalmaster, M. V. 1987. The Bald Eagle. Universe, New York.

Steenhof, K., M. N. Kochert, and T. L. Mcdonald. 1997. Interactive Effects of Prey and Weather on Golden Eagle Reproduction. Journal of Animal Ecology 66:350-362.

Steidl, R. J., and R. G. Anthony. 2000. Experimental effects of human activity on breeding bald eagles. Ecological Applications 10:258-268.

Suring, L. H. 2008. Habitat Relationships of Bald Eagles in Alaska. *in* B. A. Wright, and P. F. Schempf, editors. Bald Eagles in Alaska.

Thompson, W. L., S. Hall, and C. R. Lindsay. 2009. Evaluation of a Survey Method for Estimating Number and Monitoring Occupancy of Bald Eagle Nests in Kenai Fjords National Park. Natural Resource Technical Report NPS/SWAN/NRTR—2009/271., Fort Collins, Colorado.

U.S. Fish and Wildlife Service. 2009. Post-delisting Monitoring Plan for the Bald Eagle (*Haliaeetus leucocephalus*) in the Contiguous 48 States. U.S. Fish and Wildlife Service, Divisions of Endangered Species and Migratory Birds and State Programs, Midwest Regional Office, Twin Cities, Minnesota. 75 pp.

_____. 2015. Federal Laws to Protect Bald Eagles. <https://www.fws.gov/midwest/eagle/protect/laws.html>. Accessed 2018.

US Army Corps of Engineers, O. D. 2008. Missouri River Recovery Program Bald eagles and cottonwood trees: a dynamic duo. Paper 21.

Wilson, T. L., J. H. Schmidt, B. A. Mangipane, R. Kolstrom, and K. K. Bartz. 2018. Nest use dynamics of an undisturbed population of bald eagles. Ecology and Evolution 8:7346-7354.

Wilson, T. L., E. A. Weiss, T. Shepherd, L. M. Phillips, and B. Mangipane. 2016. Monitoring Bald Eagles in Southwest Alaska: Protocol and Standard Operating Procedures. Natural Resource Technical Report NPS/SWAN/NRR—2016, National Park Service, Fort Collins, Colorado.

Wilson, T. L., E. A. Weiss, T. Shepherd, L. M. Phillips, and B. Mangipane. 2017. Monitoring Bald Eagles in Southwest Alaska Network Parks: Protocol Narrative. NPS/SWAN/NRR—2017/1382, Natural Resource Report, Fort Collins, Colorado.

FIGURES AND TABLES

Figure 1.1. This map shows the breeding, non-breeding, and year-round ranges of bald eagles in North America (Buehler 2000).

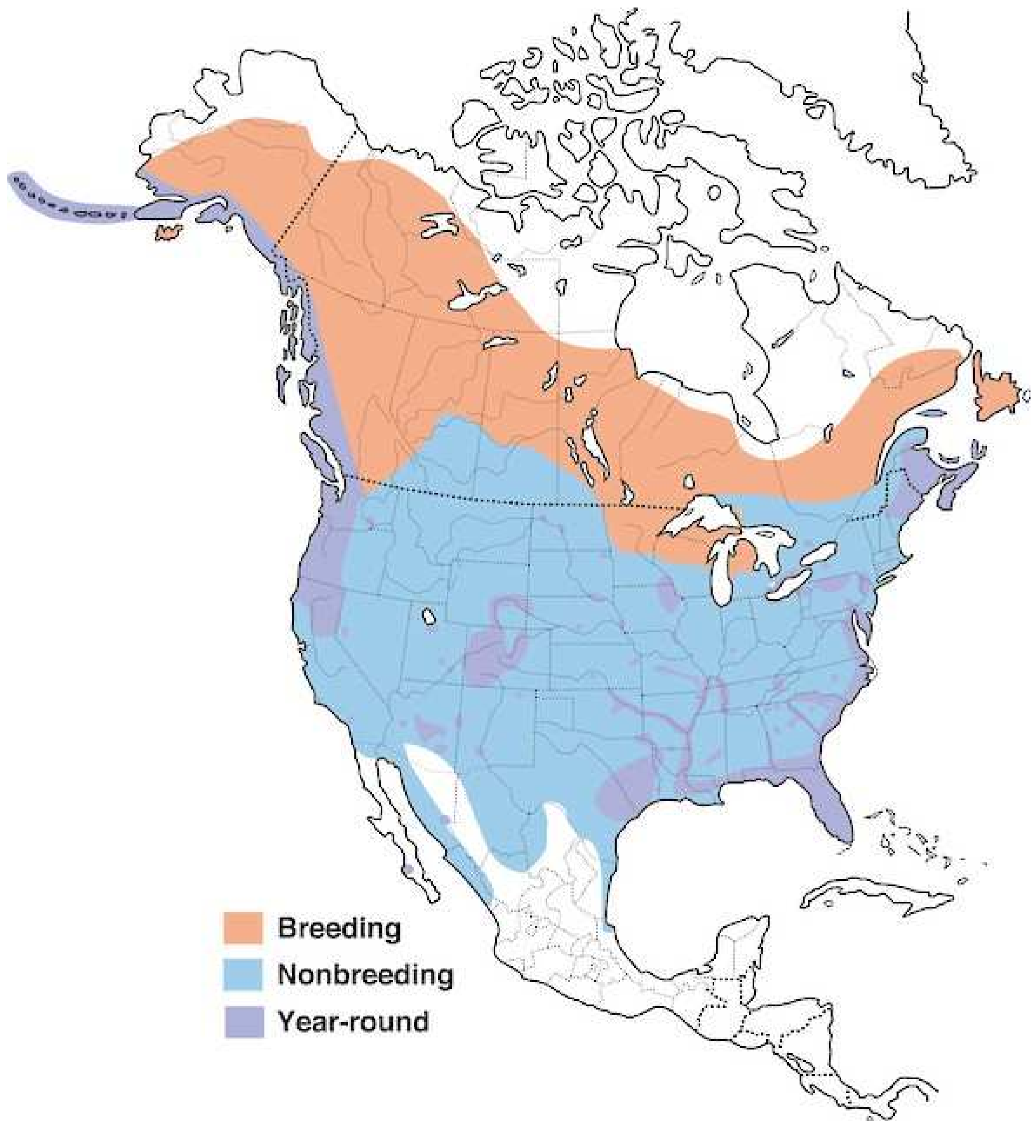


Figure 1.2. This map shows the four parks included in this study. Lake Clark National Park and Preserve, Katmai National Park and Preserve, and Kenai Fjords National Park are part of the Southwest Alaska Inventory and Monitoring Network. Wrangell – St. Elias National Park and Preserve is part of the Central Alaska Inventory and Monitoring Network.

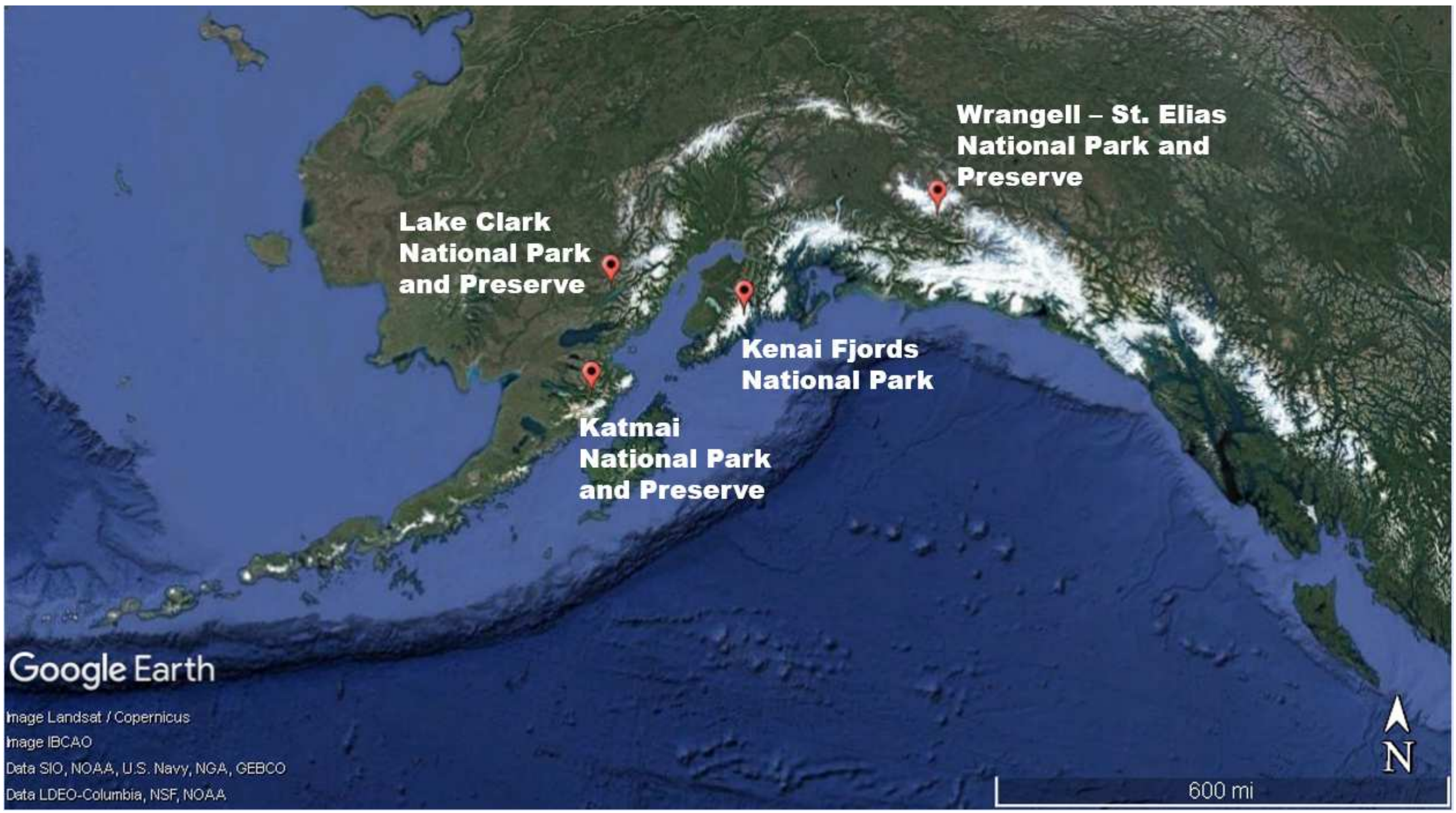


Figure 1.3. This figure is modified from the paper "A road map for designing and implementing a biological monitoring program" (Reynolds et al. 2016). It outlines the main steps that should be taken when initiating a monitoring program. Several key steps were missed when designing the bald eagle monitoring program for SWAN parks.

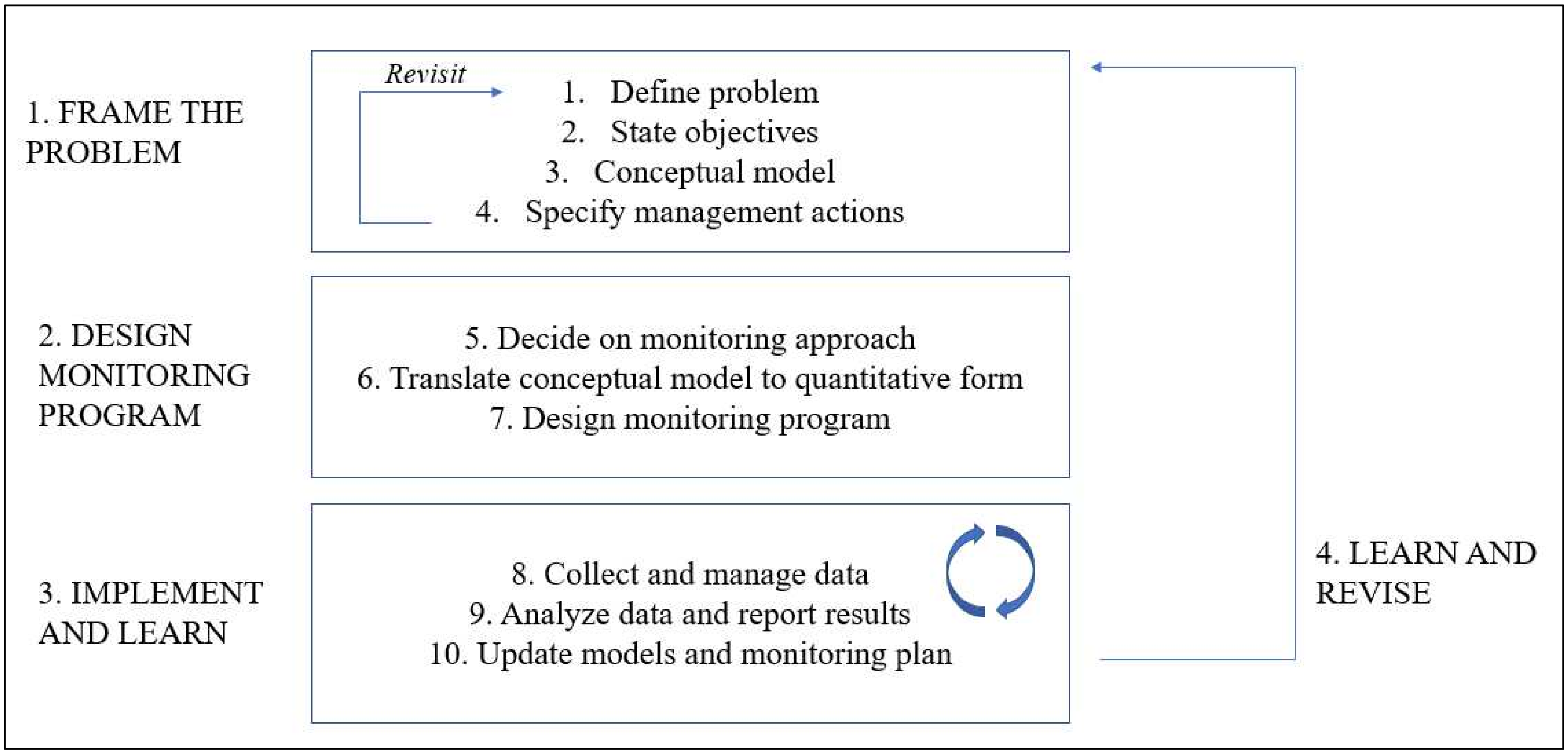


Table 1.1. This chart displays the survey methods used by Kenai Fjords National Park (KEFJ), Katmai National Park and Preserve (KATM), Lake Clark National Park and Preserve (LACL), and Wrangell – St. Elias National Park and Preserve (WRST) for bald eagle surveys. This includes sample frame, observer method, and survey period. This table is adapted from Wilson et al. (2016) to include information about WRST from Putera and Miller (2018).

	Sample Frame				
Park	**List Frame**	**Area Frame**	**Double Observer**	**Target Metric**	**Survey Period**
KEFJ	x	x	x	Nest	2009 – present
KATM	x	x	x	Nest	2010 – present
LACL	x			Nest	1992 – present
WRST	x			Territory	1987 – 1997; 2003 – present

CHAPTER 2

USING THE DELPHI PROCESS TO GATHER INFORMATION FROM A BALD EAGLE EXPERT PANEL

ABSTRACT

Bald eagle (*Haliaeetus leucocephalus*) populations are classified by the Southwest Alaska Network (SWAN) of the National Park Service as a vital sign of biological integrity, largely because of their importance as an indicator species for environmental contaminants and human disturbance. Though bald eagles are plentiful in Alaska, it is still imperative to have a monitoring plan that allows for the estimation of population sizes and detection of significant changes in populations. Currently, bald eagles are monitored in Kenai Fjords National Park, Katmai National Park and Preserve, Lake Clark National Park and Preserve, and Wrangell – St. Elias National Park, but each park uses different monitoring procedures and evaluation criteria. This makes it difficult for scientists and managers to compare data, detect changes in overall populations, and make effective management decisions. Our research is using a formal structured decision making process to ensure that the bald eagle monitoring conducted by the parks is standardized and meets programmatic goals and objectives. We implemented a Delphi process, which is an iterative survey technique that is used to gather expert opinion. We used online questionnaires to gather information and opinions from stakeholders, including National Park Service scientists and managers, eagle experts, and other interested parties. We identified important stressors and feasible monitoring metrics, which were tied to the fundamental objectives for the bald eagle monitoring program: minimize cost, minimize effort, maximize ability to detect change in populations, and

maximize accurate information about bald eagles. We will also analyze monitoring metrics using a consequence table, which determines the performance of each objective in terms of the fundamental objectives chosen by expert panelists. This information will help to create a more accurate conceptual model of the system and will eventually lead to an optimal bald eagle monitoring program that can be standardized among Southwest Alaska National Parks.

INTRODUCTION

Bald eagles are abundant in Alaska, with populations in the state estimated around 30,000 (Alaska Department of Fish and Game 2017). Southwest Alaska provides suitable coastal habitat for bald eagles, many of which reside on National Park Service land in this area (Wilson et al. 2017). The Southwest Alaska Inventory and Monitoring Network (SWAN) is comprised of five units of the National Park Service, including coastal parks Katmai National Park and Preserve (KATM), Kenai Fjords National Park (KEFJ), and Lake Clark National Park and Preserve (LACL) (Bennett et al. 2006, National Park Service 2018b). Along with Wrangell – St. Elias National Park and Preserve (WRST), which is part of the Central Alaska Inventory and Monitoring Network, these parks are home to large populations of breeding bald eagles (Wilson et al. 2017, National Park Service 2018c). Bald eagles in the parks are monitored annually by SWAN as part of their Vital Signs Monitoring Plan, which specifies the definition of vital signs monitoring as “the collection and analysis of repeated observations or measurements to evaluate ecological changes in the condition of park resources” (Bennett et al. 2006, Wilson et al. 2017).

In the process of creating a bald eagle monitoring program for the Southwest Alaska Inventory and Monitoring Network, decision-makers missed key portions of framing the problem and designing objectives for a decision about monitoring (Reynolds et al. 2016). As a result, the parks currently collect data on bald eagles slightly differently from one another and are not able to use their data as effectively as possible. Along with a screening estimator to generate population estimates, the parks use variations on the double-observer method (to account for imperfect detection of nests in area-frame counts) and dual-frame method (to provide a more rigorous count of total nests by estimating detection probabilities specific to each observer) (Haines and Pollock 1998, Nichols et al. 2000, Alpizar-Jara et al. 2005, U.S. Fish and Wildlife Service 2009).

Kenai Fjords National Park has been conducting aerial surveys along the coast since 2009 using a double observer and dual frame method. Katmai National Park and Preserve has been surveying for bald eagles in the entire Naknek watershed since 2010, also using the double observer and modified dual frame method. Lake Clark National Park and Preserve, however, uses only a single observer and the list frame to conduct their surveys, but they have collected an abundance of data from all potential bald eagle habitat in the park since 1992 (Wilson et al. 2017). Wrangell – St. Elias National Park and Preserve monitors nesting bald eagle territories annually along the Copper River (National Park Service 2018a). This discrepancy between monitoring methods makes it more difficult for park managers to use bald eagles as indicators of environmental quality through wider population trends. Through flight surveys, SWAN parks currently monitor the following metrics: “mean number of pre-fledging chicks produced in nests with spring nesting activity”, “proportion of nests that are successful at producing at least one pre-fledging chick", "total number of bald eagle nests", "total number of nesting pairs", and "proportion of nests that are used by bald eagles for reproduction" (Wilson et al. 2017). It is also important to note that, unlike SWAN parks, Wrangell – St. Elias National Park and Preserve calculates some of these metrics using bald eagle nesting territories, rather than individual nests. This inconsistency makes it difficult to examine and compare information between WRST and SWAN parks.

Despite the wealth of bald eagle monitoring data already collected in the parks, the scientists and managers cannot use data to make comparisons between parks and recognize larger, region-wide trends in populations. Bald eagle populations in the parks are relatively stable (Wilson et al. 2018), however there are various factors that affect bald eagles in these areas. According to expert panelists, the availability of salmon as a food source, suitable nesting sites, seasonal changes in prey species, potential oil spills, contaminants such as mercury, lead, and PCBs, suitability of weather conditions, and disturbance by humans may all have effects on bald eagles in Southwest Alaska National Parks. Since it is unlikely that there will be significant biological impacts to bald eagles as a result of most stressors, park managers must make decisions that balance both the bald eagle populations and park values (Gende et al. 2018, Wilson et al. 2018). It is important for parks to come to a consensus on appropriate monitoring metrics, timing of surveys, and the objectives that will underlie the bald eagle monitoring program.

Since bald eagle monitoring is ongoing and part of the vital signs monitoring program, it is not reasonable to postpone management or monitoring changes until experimental changes in monitoring have been explored (Fackler et al. 2014). Instead, an adaptive approach, that aims to reduce scientific uncertainty in dynamic systems, will allow decision-makers to evaluate tradeoffs and use the current knowledge to make the monitoring program useful for making management decisions. An adaptive monitoring program should be based on a conceptual model of the system and should focus on addressing specific and well-defined questions (Lindenmayer and Likens 2009). The Delphi process will help to gather information about bald eagles in the Southwest Alaska system and the bald eagle monitoring program, so that managers can make a decision about how to best monitor bald eagles in these protected areas (Lyons et al. 2008, Fackler et al. 2014).

Structured decision making is a transparent way to confront complex problems with multiple objectives, using group decision-making techniques to generate alternatives (Gregory et al. 2012). It provides a formal way for stakeholders to be involved in the decision-making process (Wilson and Arvai 2011). Structured decision making can be achieved using many different techniques, and value-focused structured approaches have been shown to urge decision-makers to make more informed and better-quality decisions (Arvai et al. 2001). We employed the Delphi method for this project and used intensive surveys of expert opinion to create a more complete conceptual model of the system and gather information to create a consequence table for potential monitoring metrics, that will help parks to come to a reliable consensus on the most appropriate monitoring program for bald eagles in Southwest Alaska National Parks (Dalkey and Helmer 1963).

Originating at the end of the 1940’s as part of a classified military defense project and named after the oracle at Delphi, the Delphi Process facilitates collecting expert opinion to aid in decision making when there is insufficient scientific information or contradicting information and opinions (Hasson et al. 2000, Landeta 2006). Based on a flexible methodology, and premise that a collective group belief is more reliable than that of a single expert, the Delphi process consists of a systematic series of questionnaires to guide a group of experts toward a decision or consensus opinion. The process includes controlled feedback, provided to panelists between rounds of questioning, which gives them the opportunity to consider the responses of other experts and change their own responses (Rowe and Wright 1999, Skulmoski et al. 2007, Steurer 2011). Along with this controlled feedback, essential features of a Delphi Process are anonymity of panelists (at least in their responses) to reduce negative elements of group interaction, iteration of questions to encourage critical thinking, and statistical summaries of group response (Rowe and Wright 1999, Steurer 2011).

A simple, graphical representation of a typical Delphi process is found in Figure 2.1 and adapted from Heuer and Pherson (2011). The facilitator sends the expert panel a series of structured questionnaires. Between each round of questioning, the experts receive a summarized report of panelist responses. They are permitted to change their own responses after reviewing the summary, and before the next questionnaire. This process continues until reasonable consensus is reached, or a sufficient amount of information has been collected (Skulmoski et al. 2007). The goal of the Delphi Process is not only consensus among panelists or sufficient information collected, but also to highlight dissenting viewpoints and opinions, since acknowledgment of disagreement leads to overall better-quality outcomes (Priem and Harrison 1995, Skulmoski et al. 2007, Steurer 2011).

Additionally, there are several constraints on this project that make the Delphi process a favorable option. The expert panel consists of scientists and managers located throughout Alaska and in the contiguous United States. Frequent, face-to-face meetings are not feasible, due to time and cost limitations. The Delphi process can be conducted online and facilitates the communication process among experts who are geographically distant and may have limited time (Linstone and Turoff 2002, Geist 2010). Occasional in-person communication helps to offset the criticism that the Delphi Process lacks the ability to create natural and spontaneous conversation (de Loe 1995).

The objectives for this study are to 1) Compile a comprehensive panel of experts who can provide reliable information about bald eagle monitoring in Southwest Alaska National Parks; 2) Gather information through a Delphi Process, including: important stressors that affect bald eagles, reliable and achievable monitoring metrics, cost and effort required to carry out those monitoring metrics, and fundamental objectives for the bald eagle monitoring program in Southwest Alaska National Parks; 3) Link stressors to responsive and feasible monitoring metrics, and link those metrics to the fundamental objectives for the monitoring program to create an influence diagram of the monitoring scenario. 4) Analyze monitoring metrics using a consequence table, to evaluate the performance of each objective in terms of the specified fundamental objectives for the bald eagle monitoring program. By achieving these objectives, we will have compiled information to form a more complete picture of the bald eagles in the Southwest Alaska system and bald eagle monitoring program in Southwest Alaska National Parks. This will eventually lead to the optimization and standardization of bald eagle monitoring among these National Parks. Our approach demonstrates a mechanism which ensures that monitoring is “meaningful” and can inform science-based decision making, as envisioned by Oakley et al. (2003).

METHODS

After defining the problem, we selected an expert panel to respond to questionnaires and provide input, which we refer to as the "Bald Eagle Expert Panel." As its name implies, the panel included eagle and raptor experts, as well as scientists and managers. These panelists were selected through purposive sampling, since their knowledge can be applied to this specific problem (Hasson et al. 2000). Many of these participants are involved with the Southwest Alaska Inventory and Monitoring Network. Additional participants were recruited to participate from the National Park Service, U.S. Fish & Wildlife Service, and South Dakota Game, Fish & Parks. The 18 panelists were gathered using a "snowball process." This means that we generated an initial list of interested participants and gave them the task of suggesting additional experts who may be qualified or interested in participating (Eycott et al. 2011). We then contacted these new individuals, gave them background information about the process, and asked them to participate. This continued until no new names were suggested.

We distributed all questionnaires to panelists via email. We sent Questionnaire 1 as a Microsoft Word document, which panelists completed and returned via email. We created subsequent questionnaires using SurveyMonkey. We gave panelists one month from the date the questionnaire was sent to complete the survey. Panelists received several reminder emails to complete each questionnaire. After receiving questionnaire summaries, we gave panelists two weeks to contact the facilitator with any changes they would like to make to their responses. These deadlines were flexible, and we gave panelists additional time to complete questionnaires if it was requested. Additionally, we conducted an in-person meeting on March 29, 2018 in Anchorage Alaska (in the time frame between Questionnaire 3 and Questionnaire 4), where we formulated fundamental objectives for a future analysis of monitoring scenarios.

Themes of each questionnaire in this process are found in Figure 2.2. We designed Questionnaire 1 as a tool to gather preliminary information and focus the panelists on the overall themes that would be covered during the Delphi Process. Questionnaire 1 addressed several broad themes to gather important information to form a conceptual model of bald eagles in Southwest Alaska National Parks. As focuses in this questionnaire, we included: current management and potential management of bald eagles, importance of long-term bald eagle monitoring, and stressors that impact bald eagles. We asked panelists to list various factors that affect bald eagle populations in their management areas – this formed the comprehensive list of 18 stressors, that was used to develop subsequent questionnaires. We asked questions primarily in an open-ended manner, as not to stifle any thoughts or opinions of panelists in the early stages of the process.

The summary of Questionnaire 1 largely involved qualitative data analysis. We manually completed this analysis by identifying broad themes and categorizing panelist responses. We displayed data graphically for several questions, using proportions of panelists whose responses corresponded to specific themes.

We designed Questionnaire 2 to build on the responses received from panelists in Questionnaire 1. The overall goals for this questionnaire were to elicit feedback on the importance of the 18 stressors on bald eagles (identified in Questionnaire 1) and to explore current monitoring and management for bald eagles in Alaska National Parks. Panelists ranked broad categories of stressors (created through qualitative data analysis of

Questionnaire 1 responses) based on impact to bald eagles. Five categories were presented to them, and panelists were asked to rank these categories from most impact to least impact on bald eagles. Expert panelists also rated each of the 18 specific stressors on a 5-point scale from “Not important” to “Important.”

For the Questionnaire 2 summary, we again used manual qualitative data analysis to distill open-ended panelist responses. To analyze the ranking of the five, broad categories of stressors and ratings of importance, we used weighted averages (Formula 1). A ranking of 1 (most impactful) received the highest value of 5. A ranking of 2 received a value of 4. A ranking of 3 received a value of 3. A ranking of 4 received a value of 2, and a ranking of 5 received the lowest value of 1. Categories with higher weighted averages were evaluated as having more impact on bald eagles.

Formula 1. *This formula is used to calculate weighted averages for rating and multiple choice questions.*

$$\frac{x_1 w_1 + x_2 w_2 + x_3 w_3 \ldots x_n w_n}{Total},$$

$$where\ w = weight\ of\ ranked\ position, x = response\ count\ for\ answer\ choice$$

We designed Questionnaire 3 to further examine important stressors to bald eagles in the system, and to begin to address monitoring metrics. An additional goal for this questionnaire was to quantify a link between stressors and monitoring metrics. Panelists were asked to re-evaluate the most important stressors to bald eagles by ranking their top five stressors. Next, experts were instructed to rate the reliability of 15 monitoring metrics on a 5-point scale from “Extremely Unreliable” to “Extremely Reliable.” They were also asked to provide additional monitoring metrics that were not included in the list. Panelists were then asked to select all monitoring metrics (of the original 15 metrics provided) that would respond to changes in each of the 18 stressors.

Productivity is an imperfect surrogate of true population demographics, because it is impossible to verify nest initiation or successful fledging using aerial surveys. It is an inherently flawed demographic index that combines elements of fecundity and recruitment. There are several methods used to obtain this metric, all of which use the term productivity. This has caused substantial confusion in the literature (McIntyre and Schmidt 2012). A true population demographic would require an accurate measure of how many eggs were laid, and how many of those eggs successfully hatch and are recruited into the population. Four of our 15 metrics were designed to monitor various aspects of productivity. “Mean number of chicks fledged per nesting pair,” “Total number of chicks produced,” “Mean number of young produced per nesting pair,” and “Mean number of pre-fledging chicks produced in nests with spring nesting activity” were designed to describe similar, but different ways to measure of productivity. By describing the metrics differently, we hoped to emphasize that, under the current protocol, absolute numbers of eggs laid or chicks fledged are not confirmed without additional survey effort. Rather than use just one metric, “Productivity”, our different versions of this metric attempted to query our experts about how much uncertainty they are willing to accept in these measures of true population demographic parameters.

However, we did not clearly communicate our intentions with the panel, causing considerable confusion about what was meant by the subtle differences in each metric description. In subsequence, we will combine these four metrics into a single metric, “Productivity,” instead of addressing each metric separately.

Once again, we used manual qualitative data analysis to analyze open-ended responses. We analyzed the ranking of the most important stressors using weighted averages (Formula 1). We assigned values to each ranking position so that stressors with higher weighted averages are ranked as more important. When ranking reliability, we assigned a value to each multiple choice response: 1 = Extremely Unreliable; 2 = Not Reliable; 3 = Somewhat Reliable; 4 = Reliable; 5 = Extremely Reliable. We calculated weighted averages for each metric and considered metrics with higher weighted averages to be more reliable. For questions asking panelists which metrics are responsive to changes in stressors, panelists were permitted to choose more than one response. Therefore, we used frequency distributions to display results to panelists.

The purpose of Questionnaire 4 was to gather information about monitoring metrics suggested by panelists in Questionnaire 3. We asked panelists questions about reliability and responsiveness of these new metrics, so that they could be combined with the metrics from Questionnaire 3. We also asked questions about the cost of conducting one year of monitoring for each metric. Panelists were given multiple choice options with cost ranges from $0-5,000 to $25,000+, and were asked to select the most appropriate response. As an additional metric of cost, we asked experts to estimate the number of annual person days each monitoring metric would require.

As with previous questionnaires, we used manual qualitative analysis for open-ended response. We calculated reliability and responsiveness of monitoring metrics the same way as for the original 15 monitoring metrics in Questionnaire 3. We assigned values to multiple choice responses for questions regarding cost: 1 = $0-5,000; 2 = $5,000-10,000; 3 = $10,000-15,000; 4 = $15,000-20,000; 5 = $20,000-25,000; 6 =

$25,000+. We calculated a weighted average for each metric, and metrics with higher weighted averages are valued as more expensive than those with lower weighted averages. We displayed effort to panelists using boxplots, and also calculated mean annual person days for each metric.

An additional component to this process was an in-person meeting with all available panelists. Attendees included those who were able to travel to Anchorage, AK on March 29, 2018, or video conference into the meeting. This meeting involved a presentation of the results collected to that point, as well as several hours for panelist discussion. We covered several topics of discussion, including current needs and desires for the bald eagle monitoring program, generating fundamental objectives for the bald eagle monitoring program, and the metrics that might be used to measure success in those fundamental objectives.

We synthesized information collected from questionnaires in several ways. First, we analyzed monitoring metrics based on their composition of reliability score (weighted average), cost score (weighted average), and estimate of effort (mean annual person days) to determine feasibility. We created a weighted scatterplot with cost on the x-axis, reliability on the y-axis, and effort designating the size of each point, and divided it into four sections based on cost and reliability.

We divided sections based on weighted average thresholds that correspond to multiple choice responses. We placed the threshold for Reliable vs. Unreliable at a weighted average value of 3.0. This value of 3 corresponds to the multiple choice response: Somewhat Reliable. We drew a horizontal line on the weighted scatterplot to denote the reliability threshold. We placed the threshold for determining Low Cost vs.

High Cost at the center of the range of weighted averages assigned to metrics by panelists (weighted averages of cost ranged from 3.0-6.0). Therefore, we placed the cost threshold at a level of 4.5. We drew a horizontal line on the weighted scatterplot to denote the cost threshold. We included mean effort in this compilation by altering the weights of each point in the weighted scatterplot. Unlike cost and reliability, we will not use effort specifically to exclude metrics, but it will be considered in any decisions that are made.

We then combined important stressors and feasible monitoring metrics into an influence diagram, which incorporates the fundamental objectives generated by panelists in the in-person meeting. We included all four of the fundamental objectives in the influence diagram, based on information collected in the questionnaires. This influence diagram connects important stressors to bald eagles to feasible monitoring metrics. In this diagram, we used only the seven most important stressors to bald eagles, as identified by panelists. We included only monitoring metrics that were placed in the "Low Cost, Low Reliability" and "Low Cost, High Reliability" sections. The four fundamental objectives: "Minimize Cost," "Maximize Accurate Information about Bald Eagles," Minimize Effort," and "Maximize Ability to Detect Change" were included in the diagram.

We created links between stressors and metrics based on data from panelists in Questionnaire 3 and Questionnaire 4, which identified monitoring metrics that will respond to changes in each of the stressors. A review of the definition of consensus in 100 Delphi studies showed that for processes using a percentage of panelists to define consensus, 75% was the median threshold (Diamond et al. 2014). Using this threshold, we drew links between each stressor and metrics that were identified by ≥75% of panelists as being responsive to that stressor. We then connected monitoring objectives to relevant fundamental objectives. For the objective “Minimize Cost,” we drew links from monitoring metrics that fell in the “Low Cost, Low Reliability” and “Low Cost, High Reliability” categories. We linked the objective “Minimize Effort” to monitoring metrics, based on metrics that have mean effort values of less than 20 annual person days, as estimated by panelists. We linked the objective “Maximize accurate information about bald eagles” to metrics that fell into the “Low Cost, High Reliability” category.

We linked feasible monitoring metrics with the fourth objective, “Maximize Ability to Detect Change,” by creating a contingency table (simplified contingency table shown in Table 2.3). This contingency table examines the sensitivity of feasible monitoring metrics to important stressors. Stressors are denoted in each of the rows, and feasible monitoring metrics are denoted in each of the columns. Each cell received a value of either 0 or 1. A cell at the intersection of a metric and a stressor receives a value of 1 if panelists rated that metric as responsive to the stressor. Alternatively a cell receives a value of 0 if panelists rated that metric as non-responsive to a stressor. We consider metrics to be responsive to a stressor if $\geq$75% of panelists rated it as such. The sum of each row denotes how many metrics may measure a change in each stressor. The sum of each column represents the sensitivity of each metric, or how responsive that metric is to changes in important stressors. Metrics are assigned a link to the fundamental objective “Maximize Ability to Detect Change” if that metric is responsive to >50% (a sensitivity value of four or more) of the full list of stressors to bald eagles.

We then analyzed monitoring metrics using a consequence table (Table 2.4). For this analysis, we combined four metrics: “Mean number of chicks fledged per nesting pair”, “Total number of chicks produced”, “Mean number of young produced per nesting pair", and "Mean number of pre-fledging chicks produced in nests with spring nesting activity." We renamed this new combined metric "Productivity." This was done to reduce confusion among panelists and simplify the monitoring metrics. We listed each fundamental objective vertically, along with corresponding performance measures. The full suite of monitoring metrics (including the new metric, "Productivity") are displayed horizontally. In the intersecting cells, we listed the estimated consequences for each monitoring metric, as they relate to each fundamental objective. Inputs for these estimated consequences are described, below.

Inputs for the objective "Minimize Cost" are taken from panelist responses to questions asking about the annual cost for performing each metric. The values in the consequence table are the weighted averages calculated from multiple choice responses. Inputs for the objective "Minimize Effort" are taken from the mean of panelist responses to a question asking to estimate the annual number of person days each metric would require. Inputs for the objective "Maximize Ability to Detect Change" are a measure of sensitivity. This value is a count of the stressors on bald eagles, to which each metric is responsive. These values come from panelist responses to questionnaires and the original list of 18 stressors. Inputs for the objective "Maximize Accurate Information About Bald Eagles" are the weighted averages of panelist responses about the reliability of each metric.

Since we did not directly query the panel about a combined productivity metric, to gather inputs for the "Productivity" metric, we approximated values using another metric as a reference point. Specifically, we used the metric "Proportion of nests successful at producing at least one pre-fledging chick" as a reference point. This metric requires two

May surveys, whereas productivity surveys can be achieved using only one May survey. Therefore, the cost and effort are slightly lower for "Productivity" than for "Proportion of nests successful at producing at least one pre-fledging chick," and the ability to detect change and accurate information about bald eagles will be slightly higher. We asked panelists to readjust numbers in the consequence table, that they felt did not accurately describe the performance of each monitoring metric.

Using the inputs described above, we used the consequence table to eliminate dominated metrics. If a metric performs better than another metric in all four fundamental objective categories, the lesser performing metric is outcompeted. The lesser performing metric is subsequently eliminated from consideration for the monitoring program. This simplified the list of metrics that would be considered for the monitoring program.

RESULTS

Response rates for each questionnaire are shown in Table 2.1. Submitted responses to questionnaires, as well as confirmed non-responses were factored into the "Percent Response." Because Delphi participants are permitted to be only observers if they choose, these confirmed non-responses were included in the response rate.

Qualitative data analysis of responses to the four questionnaires yielded several noteworthy points about bald eagle monitoring and management in Southwest Alaska National Parks. First, all panelists agreed on the importance of frequent bald eagle monitoring. Since bald eagles are protected legally (by the Migratory Bird Treaty Act of 1918, Lacey Act of 1900, and Bald and Golden Eagle Protection Act of 1940), are an important indicator species in SWAN parks, and hold value to many park visitors, panelists agreed that it is important to monitor and protect this species. Although panelists agree that changes in bald eagle populations are likely to be slow and gradual (barring a catastrophic event or large point source contamination), several experts noted that less frequent monitoring (currently, monitoring occurs annually in SWAN parks) will make identifying trends in populations more difficult.

Panelists identified a list of 18 stressors that influence bald eagles in their management areas. These include: availability of salmon, nest site availability/suitable nesting locations, seasonal changes in prey, oil spills, mercury, lead, and PCBs, weather conditions, human disturbance, eating lead-contaminated carcasses, increases in visitation, development, ocean acidification, loss of nests to wildfire and landslides, wind-generated power, retreat of glaciers, increase in flooding events, exotic diseases, marine debris, and illegal take of feathers. Panelists rated this extensive list of stressors by importance to come up with a simplified list of the seven, most important stressors to bald eagles living on federally protected lands in Alaska. These are: availability of salmon, nest site availability/suitable nesting locations, seasonal changes in prey, weather conditions, mercury, lead, and PCBs, oil spills, and human disturbance. Most panelists felt that this list will adequately explain the most likely causes of changes in the bald eagle population in SWAN parks, however several panelists mentioned that the possibility still exists for an unknown exotic disease or an undefined catastrophic event to impact the bald eagle population.

We began quantitative data analysis by examining the broad categories of stressors. Panelists ($n = 13$) were asked to rate the impact of broad categories of stressors that affect bald eagles in their respective management areas (Figure 2.3). Based on weighted averages calculated from panelist rankings, panelists considered food availability the most impactful category of stressors to bald eagles, followed by environmental factors, climate change, contaminants, and disturbance. Panelists rated human disturbance as having the lowest impact on bald eagles in the management areas represented by expert panelists.

All monitoring metrics generated by facilitators and panelists, as well as information collected from subsequent questionnaires (including reliability of monitoring metrics, cost of monitoring metrics, and effort of monitoring metrics) can be found in Table 2.2. We displayed reliability in terms of a weighted average of panelist response. Higher weighted averages indicate higher reliability. Based on this measure, “Adult Survival” was rated as the monitoring metric with the highest reliability and “Total number of Bald Eagle nests” was assigned the lowest reliability score. We also represented cost as a weighted average of panelist response, with higher weighted averages representing higher costs. “Subadult survival” was rated as having the highest cost, and “Total number of Bald Eagle nests” was rated as the lowest-cost monitoring metric. Effort is displayed as the mean number of annual person days that each monitoring metric would require. “Abundance of bald eagles as determined by statewide aerial surveys” is rated as requiring the most amount of effort and “Total number of bald eagle nests” was rated as having the lowest effort.

We synthesized information from Table 2.2 into a graphical representation, in Figure 2.4. We divided this graph into four categories: “Low Cost, High Reliability,” “High Cost, High Reliability,” “Low Cost, Low Reliability,” and “Low Cost, Low Reliability.” Monitoring metrics that fall into the “Low Cost, Low Reliability,” and

"Low Cost, High Reliability" categories are considered feasible monitoring metrics. Feasible metrics include: "Total number of bald eagle nests", "Changes in distribution (immigration/emigration) in the study area", "Total number of chicks produced", "Mean number of young produced per nesting pair", "Mean number of pre-fledging chicks produced in nests with spring nesting activity", "Proportion of nests used by Bald Eagles for reproduction", "Proportion of nests that are successful at producing at least one pre-fledging chick", and "Total number of nesting pairs."

The in-person meeting with available expert panelists generated more in-depth discussion and covered complex topics. The most significant information gathered from this meeting was the list of four, fundamental objectives that panelists felt should underly the monitoring program. These include: minimize the cost of the monitoring program, minimize the effort of the monitoring program, maximize the amount of accurate information collected about bald eagles, and maximize ability to detect change in bald eagle populations. Panelists also created a separate set of metrics that will be used to assess whether those fundamental objectives are being met. Cost will be measured in dollars, annually. Effort will be measured in number of annual person days. Estimates of effort will include field preparation, field surveys, and data handling. Panelists suggested that the ability to detect change should be measured in terms of minimum detectable change. As a proxy for this metric, we used the sensitivity score to determine ability to detect change. Panelists suggested that accurate information about eagles can be measured using abundance, productivity, and dynamics. As a surrogate for these more complicated metrics, we used a reliability score collected from panelists. Although not included as a fundamental objective, panelists felt that an underlying objective is the ability to integrate information about bald eagles across parks.

Table 2.3 shows a simplified contingency table, including important stressors to bald eagles and feasible metrics, as determined by graphical analysis (Figure 2.4). An influence diagram (Figure 2.5) combines the most important stressors to bald eagles in Southwest Alaska National Parks with feasible monitoring metrics, and several of the fundamental objectives outlined by panelists. Stressors and metrics are connected based on information about responsiveness, provided by panelists. Links are made between metrics and objectives based on information provided by panelists about cost, effort, reliability, and sensitivity of monitoring metrics. The metrics "Total number of bald eagle nests" and "Changes in distribution (immigration/emigration) in the study area" connect to the fundamental objective "Minimize Cost." The fundamental objective "Minimize Effort" is linked to the metrics "Total number of bald eagle nests," "Total number of nesting pairs," "Proportion of nests used by bald eagles for reproduction," and "Proportion of nests successful at producing at least one pre-fledging chick." The metrics linked to the fundamental objective "Maximize accurate information about bald eagles" are "Total number of nesting pairs," "Proportion of nests used by bald eagles for reproduction," "Total number of chicks produced," "Mean number of pre-fledging chicks produced in nests with spring nesting activity," "Mean number of young produced per nesting pair," and "Proportion of nests successful at producing at least one pre-fledging chick." The fundamental objective "Maximize Ability to Detect Change" is linked to the metrics "Total number of chicks produced," "Mean number of young produced per nesting pair," and "Proportion of nests successful at producing at least one pre-fledging chick."

The consequence table yielded compiled information from the Delphi survey. The best performing metric for the objective "Minimize cost" (Total # bald eagle nests) has a weighted average cost value of 3. The worst performing metric in this category (Abundance of bald eagles as determined by statewide aerial surveys) has a weighted average cost value of 6. The best performing metric for the objective "Minimize Effort" (Total # bald eagle nests) has a mean annual effort value of 11.5 days. The worst performing metric in this category (Abundance of bald eagles as determined by statewide aerial surveys) has a mean annual effort value of 63.3 days. The best performing metric for the objective "Maximize ability to detect change" (Changes in distribution) has a sensitivity value of 9. The worst performing metrics in this category (Abundance at winter concentration sites; Proportion of breeding to non-breeding eagles) have sensitivity values of 0. The best performing metric for the objective "Maximize accurate information about bald eagles" (Adult survival) has a weighted average reliability value of 4. The worst performing metric in this category (Total # bald eagle nests) has a weighted average reliability value of 2.27.

We used the consequence table to generate a list of monitoring metrics that remain, after elimination of practically dominated metrics. The metrics that withstood the elimination process and remain in consideration for the monitoring program include "Total number of bald eagle nests", "Changes in distribution", "Productivity", "Proportion of nests used by bald eagles for reproduction", "Total number of nesting pairs", and "Adult survival."

Our graphical analysis of monitoring metrics and the analysis of metrics using the consequence table yielded similar lists of feasible monitoring metrics. Both analyses identified “Total number of bald eagle nests”, “Changes in distribution”, “Proportion of bald eagle nests used for reproduction”, and “Total number of nesting pairs” as feasible metrics. Additionally, three of the four metrics that were combined to form “Productivity” were considered feasible in the graphical analysis, and “Productivity” was considered feasible through the consequence table analysis. There were only two inconsistencies between these analyses. The graphical analysis of metrics identified “Proportion of nests that are successful at producing at least one pre-fledging chick” as a feasible metric, while the consequence table showed this metric outcompeted by the “Productivity” metric. However, these metrics are similar and the information for “Proportion of nests that are successful at producing at least one pre-fledging chick” is completely contained in the information for “Productivity.” The consequence table identified “Adult survival” as a feasible metric, while the graphical analysis eliminated this metric based on cost.

DISCUSSION

Although National Parks in Southwest Alaska do not currently specifically manage bald eagle populations, scientists and managers should be aware of and agree on the actions that will be taken should the population experience a decline. Decision makers must make difficult choices about which subset of monitoring metrics will provide the best information about bald eagle populations, given the constraints of the program. Decision-makers will choose the optimal monitoring scenario, which balances the fundamental objectives of minimizing cost and effort, maximizing the information collected, and maximizing the ability to detect change. We used the Delphi Process to gather information from a panel of experts about the stressors that affect bald eagles, the metrics that may be used to monitor bald eagle populations, and the fundamental objectives that underlie the long-term monitoring program. By linking this information, we can better conceptualize the system, and will help to evaluate alternatives to ultimately choose an optimal monitoring program that can be standardized among Southwest Alaska National Parks.

Using the Delphi Process, we gathered information that will be used to optimize and standardize the bald eagle monitoring program in Southwest Alaska National Parks. We collected information about stressors that affect bald eagles, as well as the monitoring metrics that may be used to gain a more complete picture of the system. The themes outlined through qualitative data analysis highlight the commitment of panelists to forming an optimal bald eagle monitoring program that will allow SWAN parks to detect changes in bald eagle populations in a timely fashion. Bald eagles are subject to many stressors and the responses to these questionnaires emphasized the need to clarify how a long-term monitoring program would be responsive to changes likely to affect their populations.

Consistent with the findings that prey availability (especially early in the breeding season) influences reproductive success of bald eagles in Alaska, panelists rated food availability as the most impactful, broad influence on bald eagle populations (Hansen and Hodges 1985, Steidl et al. 1997). Despite the fact that visitation is an extremely common impact to protected areas (Gende et al. 2018), panelists rated human disturbance as the least important broad category of stressors to bald eagles. Several panelists mentioned that visitor impact in SWAN parks is mostly localized and not likely to have real, lasting effects on bald eagle populations in the parks. However, distancing nests from human disturbance using buffer zones has been shown to influence population-level responses such as increased nest occupancy and greater reproductive success (Cruz et al. 2018).

Panelists were asked to consider a large list of monitoring metrics. This list included the metrics that are currently monitored: “mean number of pre-fledging chicks produced in nests with spring nesting activity”, “proportion of nests that are successful at producing at least one pre-fledging chick”, “total number of bald eagle nests”, “total number of nesting pairs”, and “proportion of nests that are used by bald eagles for reproduction”. It is important to note that these currently monitored metrics do not completely correlate with the most reliable metrics, as rated by panelists. In fact, one of these currently monitored metrics, “total number of bald eagle nests” was rated by panelists as the least reliable monitoring metric. This comparison of currently monitored metrics with most reliable metrics accentuates the need to re-evaluate the current monitoring plan, to ensure that scientists and managers are gaining an accurate and complete picture of the status of bald eagle populations in the parks. It is also important to note that several of the most reliable metrics were also rated as the most expensive. For example, monitoring adult survival and subadult survival would be highly reliable, but since parks are limited by cost and both metrics fall into the “high cost” category, it may not be feasible for parks to begin monitoring survival of adult and subadult bald eagles. It may also be important to emphasize that all metrics in the “High Reliability, Low Cost” category are already monitored by parks, indicating that compromises may already be occurring in the monitoring program.

Metrics in both “Low Cost, High Reliability” and “Low Cost, Low Reliability” categories were considered feasible monitoring metrics. Although “Low Cost, Low Reliability” metrics are not ideal, it can be argued that for very little or minimal additional cost, it is better to collect more information on bald eagles. For example, “total number of bald eagle nests” is low in reliability but is most likely a metric that can be easily calculated during flight surveys for additional metrics. Through comparisons of cost and reliability, with consideration to effort, the list of feasible metrics includes: “Total number of bald eagle nests”, “Changes in distribution (immigration/emigration) in the study area”, “Total number of chicks produced”, “Mean number of young produced per nesting pair”, “Mean number of pre-fledging chicks produced in nests with spring nesting activity”, “Proportion of nests used by Bald Eagles for reproduction”, “Proportion of nests that are successful at producing at least one pre-fledging chick”, and “Total number of nesting pairs.” Although mean number of annual person days required for each monitoring metric was included as a measure of effort, this measure was not used to exclude metrics from consideration. This is due to small sample sizes of respondents for questions regarding effort. Future research should attempt to gain a more accurate picture of the effort needed to conduct each monitoring metric, so that this measure can be further used to determine the feasibility of metrics and exclude high-effort metrics from consideration.

An important step in structured decision making is sketching the decision to visualize the problem at hand (Gregory et al. 2012). By creating an influence diagram that connects important stressors to bald eagles in Southwest Alaska, monitoring metrics that may be used to detect changes in populations, and fundamental objectives generated by expert panelists, we can more clearly picture the complex and interconnected nature of this problem. If bald eagles experience a decline in SWAN parks that is detected by monitoring metrics, scientists and managers may be able to use this influence diagram to link the decline to possible causes and ensure that they are upholding the underlying objectives of the monitoring program.

As their names imply, the fundamental objectives of “minimize cost” and “minimize effort” were connected to monitoring metrics through information collected from panelists about the cost and effort required for each metric. The objective “maximize accurate information about bald eagles” was linked to monitoring metrics through information collected about reliability of the monitoring metrics. More reliable metrics will provide more accurate information about bald eagle populations in SWAN parks. The fundamental objective “maximize ability to detect change” was connected to monitoring metrics through the sensitivity of those metrics to changes in important stressors. It is important to note that no single monitoring metric achieves all four of the fundamental objectives that are important to panelists. Instead, a suite of metrics should be selected to gather accurate information about bald eagles and maximize the ability to detect change, while conserving money and time. Panelists must next rate fundamental objectives in order of importance, so that optimization techniques may be used to build a monitoring program.

Panelists rated “Human Disturbance” as one of the most important stressors to bald eagles in Southwest Alaska, however this stressor had no links to the feasible monitoring metrics. By human disturbance, we mean disturbance at nesting sites through anthropogenic activity. The National Park Service is responsible for both protecting resources and providing opportunities to human visitors, and though human activity may not affect the sustainability of a population, parks may still manage to uphold their values by reducing human impact (Gende et al. 2018). Human disturbance has the potential to be a significant issue to bald eagles nesting in the parks. Comments from several panelists and literature illustrate that managing for human activity does have the potential to affect breeding success of bald eagles (Cruz et al. 2018). Many studies have documented both direct and indirect effects of human activity on raptors (Richardson and Miller 1997). Grubb and King (1991) documented bald eagles flushing in response to human activity and (Stalmaster and Newman 1978) note the negative effects that humans can have on bald eagles, in both their distribution and behavior.

The consequence table allowed us to remove metrics that are not able to compete with other metrics and streamline the list of possible monitoring metrics into a smaller, more feasible list. This list of metrics matches closely to the list of currently monitored metrics, except for the metric “Proportion of nests successful at producing at least one pre-fledging chick.” This metric is currently monitored, however was outcompeted in the consequence table. This indicates that better information can be collected for less resources by a different metric, thus eliminating the need for this metric. This should be considered by SWAN, when reevaluating their monitoring program.

Sample size of respondents is one of the limiting factors in the statistical analysis of the results of the Delphi process. There is no set number of panelists that should be recruited for a typical Delphi process, but the expert panel should provide a representative sample for the specific question being asked (Steurer 2011). Boje and Murnighan (1982) found that in their study of two modified Delphi processes, increased group size did not affect the accuracy of group response in a structured setting. We felt that our bald eagle expert panel, consisting of 18 scientists and managers, was adequate for this Delphi Process. However, response rates varied throughout the process. We expect that survey fatigue contributed to a drop in the response rate, especially for more complex and detailed questions. Survey fatigue has been studied for many years, and usually, longer surveys correlate with lower response rates (Porter et al. 2004). This limited our ability to statistically analyze panelist response, and caution should be used when attempting to generalize some of the results of this study to the entire study area. It has been demonstrated that rates of drop-out may be high in later stages of Delphi processes (Day and Bobeva 2005), and we experienced this phenomenon during our Delphi Process.

Although Questionnaire 4 had a similar overall response rate to the other questionnaires, there was a significant drop in response rates for the last several questions in this questionnaire. The lack of responses may be attributed to several factors. First, these questions elicited information that was perhaps more difficult and time-consuming for panelists to generate. Second, these questions asked for specific information about Southwest Alaska National Parks. Several panelists who were not affiliated with a specific National Park in Southwest Alaska may have skipped these questions due to lack of knowledge about the specific system. Third, since panelists that are directly affected by the outcome of the Delphi process are more apt to participate and be engaged in the process, those panelists that were not affected may have chosen not to participate in these questions (Hasson et al. 2000).

Additionally, linguistic uncertainty contributed to some confusion among panelists. Ambiguity, and underspecificity in terminology used to describe productivity metrics contributed to confusion among panelists (Regan et al. 2002, McIntyre and Schmidt 2012). We were unclear in our descriptions of metrics and did not communicate our reasons for describing “productivity” several ways, which caused misunderstandings about the information being collected from the monitoring metric. Although our aim was to explore productivity metrics in various ways that would require different amounts of survey effort, it was not communicated clearly enough to our panel. Future research using these productivity measures should aim to clarify the assumptions and bias associated with various productivity metrics and population estimators.

By using the Delphi Process, we gathered information about the seven most important stressors to bald eagles on federally protected land in Alaska. We queried a panel of expert about various monitoring metrics and compiled a list of metrics that are feasible based on cost, reliability, and effort. Finally, we linked important stressors to responsive metrics, based on panelist opinion, and identified how each monitoring metric would help to achieve the fundamental objectives of this long-term bald eagle monitoring plan. Next, we will develop alternative monitoring scenarios that alter metrics used and the timing and scope of the monitoring efforts. By using a Bayesian Decision Net, we will compare each monitoring scenario, based on how well it achieves fundamental objectives and the overall utility of each monitoring scenario. This analysis will facilitate a better informed decision, that uses components of structured decision making to combine scientific knowledge and human judgement (Gregory et al. 2012).

LITERATURE CITED

Alaska Department of Fish and Game. 2017. Bald Eagle (Haliaeetus leucocephalus) Species Profile. <http://www.adfg.alaska.gov/index.cfm?adfg=baldeagle.main> Accessed 2018.

Alpizar-Jara, R., K. H. Pollock, and D. E. Haines. 2005. Mark-recapture estimators for dual frame population size of prominent nesting structures: the effect of uncertain detection probability. Environmental and Ecological Statistics 12:155-168.

Arvai, J. L., R. Gregory, and T. L. McDaniels. 2001. Testing a Structured Decision Approach Value-Focused Thinking for Deliberative Risk Communication. Risk Analysis 21:1065-1076.

Bennett, A. J., W. L. Thompson, and D. C. Mortenson. 2006. Vital Signs Monitoring Plan, Southwest Alaska Network. National Park Service, Anchorage, AK.

Boje, D. M., and J. K. Murnighan. 1982. Group Confidence Pressures in Iterative Decisions. Management Science 28:1187-1196.

Cruz, J., S. K. Windels, W. E. Thogmartin, S. M. Crimmins, L. H. Grim, B. Zuckerberg, and R. Fuller. 2018. Managing individual nests promotes population recovery of a top predator. Journal of Applied Ecology 55:1418-1429.

Dalkey, N., and O. Helmer. 1963. An Experimental Application of the Delphi Method to the Use of Experts. Management Science 9:458-467.

Day, J., and M. Bobeva. 2005. A Generic Toolkit for the Successful Management of Delphi Studies. The Electronic Journal of Business Research Methodology 3:103-116.

de Loe, R. C. 1995. Exploring complex policy questions using the policy Delphi. Applied Geography 15:53-68.

Diamond, I. R., R. C. Grant, B. M. Feldman, P. B. Pencharz, S. C. Ling, A. M. Moore, and P. W. Wales. 2014. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 67:401-409.

Eycott, A. E., M. Marzano, and K. Watts. 2011. Filling evidence gaps with expert opinion: The use of Delphi analysis in least-cost modelling of functional connectivity. Landscape and Urban Planning 103:400-409.

Fackler, P. L., K. Pacifici, J. Martin, and C. McIntyre. 2014. Efficient use of information in adaptive management with an application to managing recreation near golden eagle nesting sites. PLoS One 9:e102434.

Geist, M. R. 2010. Using the Delphi method to engage stakeholders: a comparison of two studies. Eval Program Plann 33:147-154.

Gende, S., A. N. Hendrix, and J. Schmidt. 2018. Somewhere between acceptable and sustainable: When do impacts to resources become too large in protected areas? Biological Conservation 223:138-146.

Gregory, R., L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson. 2012. Structured Decision Making: A Practical Guide to Environmental Management Choices. John Wiley & Sons, West Sussex, UK.

Grubb, T. G., and R. M. King. 1991. Assessing Human Disturbance of Breeding Bald Eagles with Classification Tree Models. The Journal of Wildlife Management 55:500-511.

Haines, D. E., and K. H. Pollock. 1998. Estimating the number of active and successful bald eagle nests an application of the dual frame method. Environmental and Ecological Statistics 5:245-256.

Hansen, A. J., and J. I. J. Hodges. 1985. High Rates of Nonbreeding Adult Bald Eagles in Southeastern Alaska. The Journal of Wildlife Management 49:454-458.

Hasson, F., S. Keeney, and H. McKenna. 2000. Research guidelines for the Delphi survey technique. Journal of Advanced Nursing 32:1008-1015.

Heuer, R. J. J., and R. H. Pherson. 2011. Structured Analytic Techniques for Intelligence Analysis. CQ Press, Washington, D.C.

Landeta, J. 2006. Current validity of the Delphi method in social sciences. Technological Forecasting and Social Change 73:467-482.

Lindenmayer, D. B., and G. E. Likens. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol Evol 24:482-486.

Linstone, H. A., and M. Turoff, editors. 2002. The Delphi Method Techniques and Applications. Murray Turoff and Harold A. Linstone.

Lyons, J. E., M. C. Runge, H. P. Laskowski, and W. L. Kendall. 2008. Monitoring in the Context of Structured Decision-Making and Adaptive Management. Journal of Wildlife Management 72:1683-1692.

McIntyre, C., and J. H. Schmidt. 2012. Ecological and environmental correlates of territory occupancy and breeding performance of migratory Golden Eagles Aquila chrysaetos in interior Alaska. The International Journal of Avian Science 154:124-135.

National Park Service. 2018a. Central Alaska Inventory and Monitoring Network: Bald Eagles. <https://www.nps.gov/im/cakn/baldeagles.htm>. Accessed 2018.

_____. 2018b. Southwest Alaska Parks and Partners. <https://www.nps.gov/im/swan/parks.htm>. Accessed 2018.

_____. 2018c. Wrangell - St. Elias National Park & Preserve Alaska: Birds. <https://www.nps.gov/wrst/learn/nature/birds.htm>. Accessed 2018.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A Double-Observer Approach for Estimating Detection Probability and Abundance from Point Counts. The Auk 117:393-408.

Oakley, K. L., L. P. Thomas, and S. G. Fancy. 2003. Guidelines for Long-Term Monitoring Protocols. Wildlife Society Bulletin (1973-2006) 31:1000-1003.

Porter, S. R., M. E. Whitcomb, and W. H. Weitzer. 2004. Multiple Surveys of Students and Survey Fatigue. New Directions for Institutional Research 2004:63-73.

Priem, R. L., and D. A. Harrison. 1995. Structured Conflict and Consensus Outcomes in Group Decision Making. Journal of Management 21:691-710.

Regan, H. M., M. Colyvan, and M. A. Burgman. 2002. A Taxonomy and Treatment of Uncertainty for Ecology and Conservation Biology. Ecological Applications 12:618-628.

Reynolds, J. H., M. G. Knutson, K. B. Newman, E. D. Silverman, and W. L. Thompson. 2016. A road map for designing and implementing a biological monitoring program. Environmental Monitoring and Assessment 188:399.

Richardson, C. T., and C. K. Miller. 1997. Recommendations for Protecting Raptors from Human Disturbance: A Review. Wildilfe Society Bulletin 25:634-638.

Rowe, G., and G. Wright. 1999. The Delphi technique as a forecasting tool: issues and analysis. International Journal of Forecasting 15:353-375.

Skulmoski, G. J., F. T. Hartman, and J. Krahn. 2007. The Delphi Method for Graduate Research. Journal of Information Technology Education 6:1-21.

Stalmaster, M. V., and J. R. Newman. 1978. Behavioral Responses of Wintering Bald Eagles to Human Activity. The Journal of Wildlife Management 42:506-513.

Steidl, R. J., K. D. Kozie, and R. G. Anthony. 1997. Reproductive Success of Bald Eagles in Interior Alaska. The Journal of Wildlife Management 61:1313-1321.

Steurer, J. 2011. The Delphi method: an efficient procedure to generate knowledge. Skeletal Radiology 40:959-961.

U.S. Fish and Wildlife Service. 2009. Post-delisting Monitoring Plan for the Bald Eagle (*Haliaeetus leucocephalus*) in the Contiguous 48 States. U.S. Fish and Wildlife Service, Divisions of Endangered Species and Migratory Birds and State Programs, Midwest Regional Office, Twin Cities, Minnesota. 75 pp.

Wilson, R. S., and J. L. Arvai. 2011. Structured Decision Making: Using decision research to improve stakeholder participation and results. Oregon State University. <https://seagrant.oregonstate.edu/sites/seagrant.oregonstate.edu/files/sgpubs/onlineonli/h11001.pdf> Accessed 2018.

Wilson, T. L., J. H. Schmidt, B. A. Mangipane, R. Kolstrom, and K. K. Bartz. 2018. Nest use dynamics of an undisturbed population of bald eagles. Ecology and Evolution 8:7346-7354.

Wilson, T. L., E. A. Weiss, T. Shepherd, L. M. Phillips, and B. Mangipane. 2017. Monitoring Bald Eagles in Southwest Alaska Network Parks: Protocol Narrative. NPS/SWAN/NRR—2017/1382, Natural Resource Report, Fort Collins, Colorado.

FIGURES AND TABLES

Figure 2.1. A representation of the Delphi Process. This figure is adapted from (Heuer and Pherson 2011). A facilitator gathers a panel of experts. These experts then receive a series of structured questionnaires, and are provided summarized feedback between each questionnaire. This process ends in reasonable consensus.

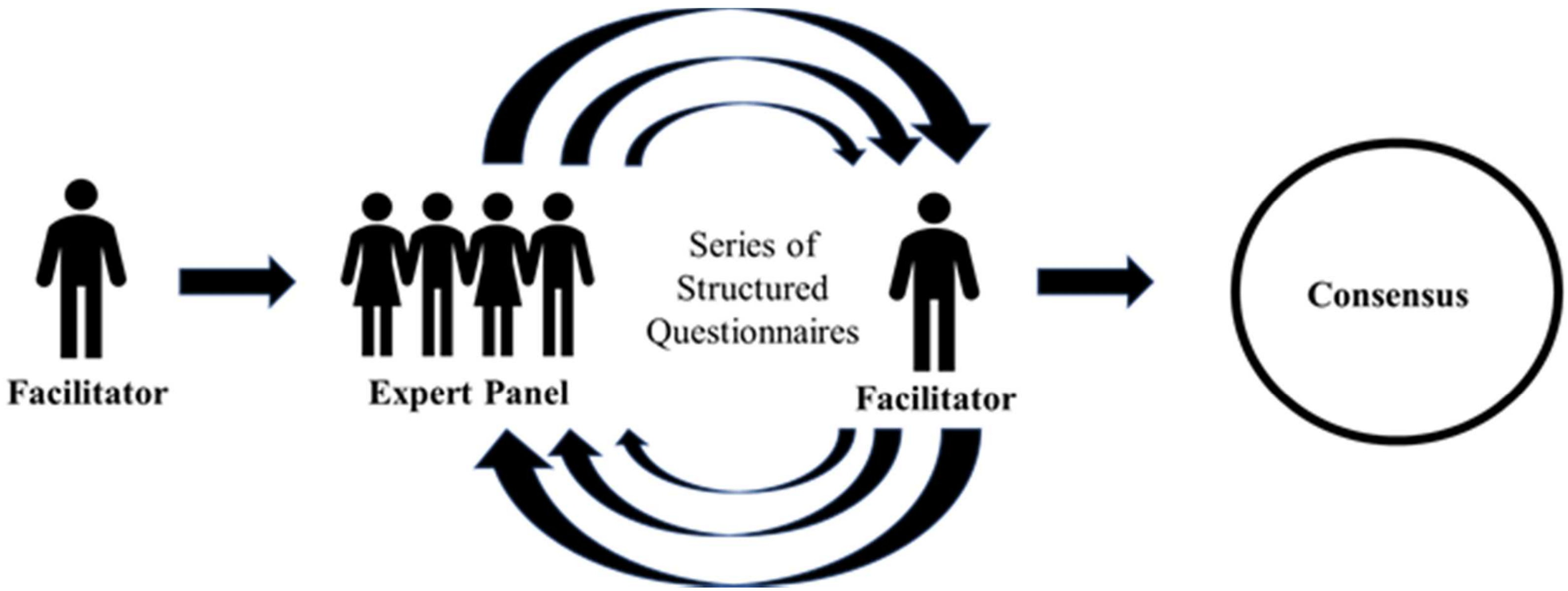


Figure 2.2. This figure outlines the general themes that will be addressed in each questionnaire round of this Delphi Process. We will use four questionnaires, with controlled feedback provided to panelists between each round of questioning.

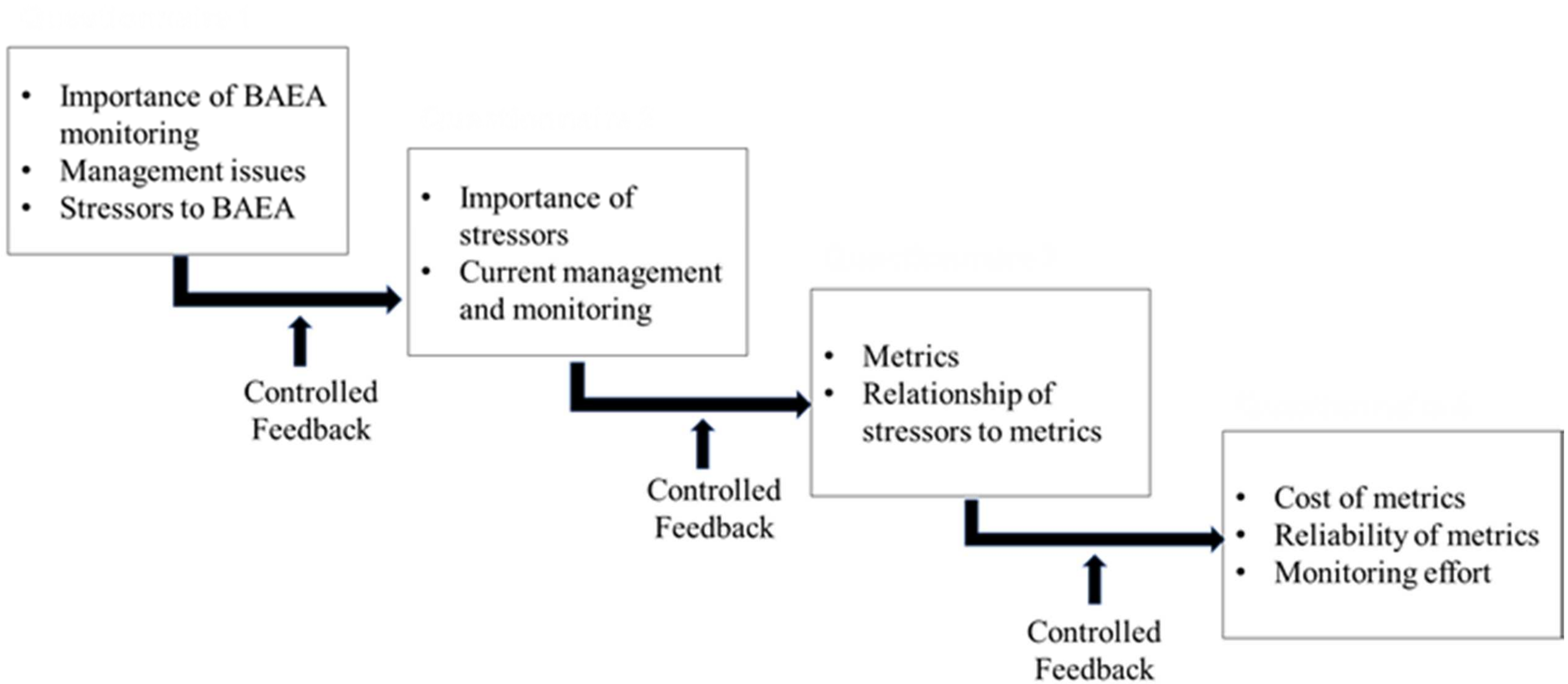


Figure 2.3. A representation of the impact of broad categories of stressors to bald eagles, indicated by weighted averages. Higher weighted averages represent categories of stressors that were rated by panelists as having more impact to bald eagles.

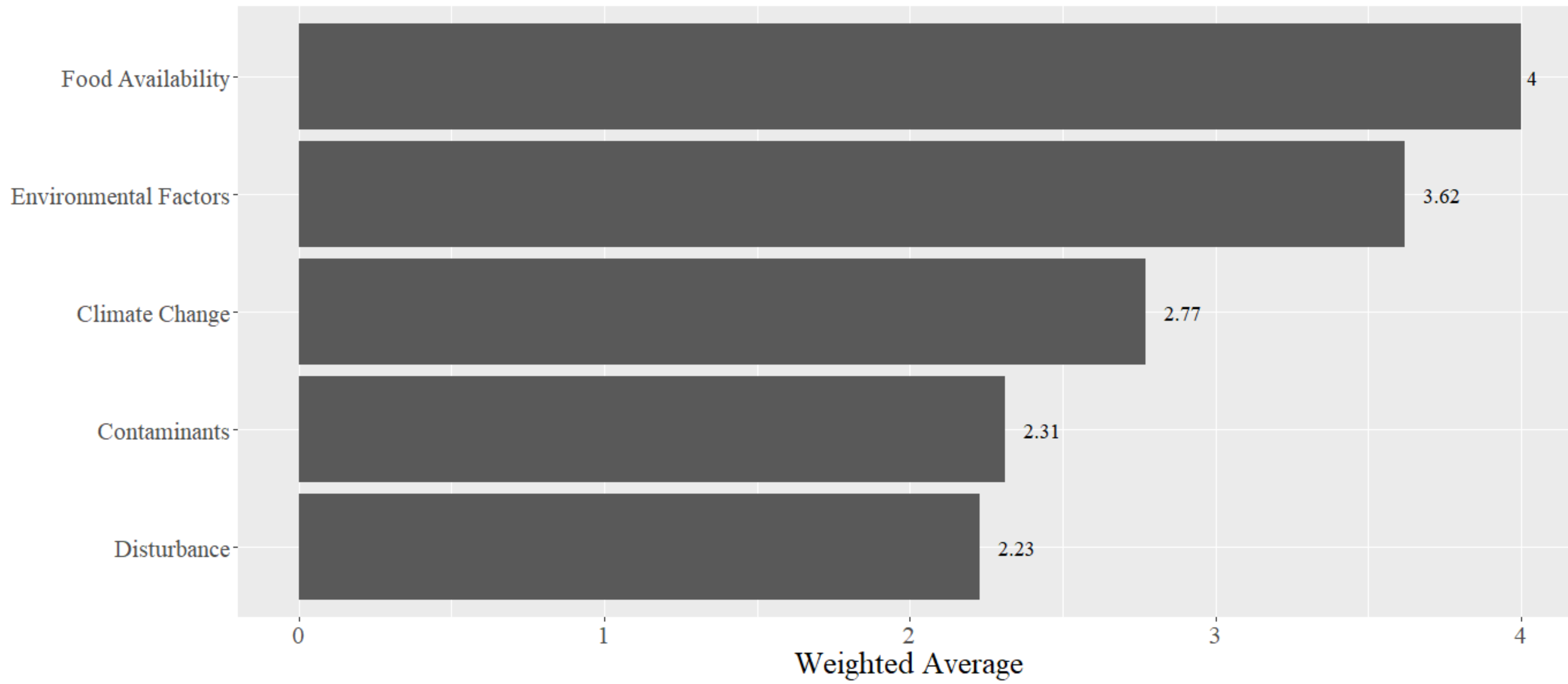


Figure 2.4. This chart categorizes monitoring metrics based on cost and reliability, and displays mean effort using a weighted scatterplot. Reference labels (see Table 2.2) are used in this graph. Monitoring metrics are divided into four categories, based on cost and reliability.

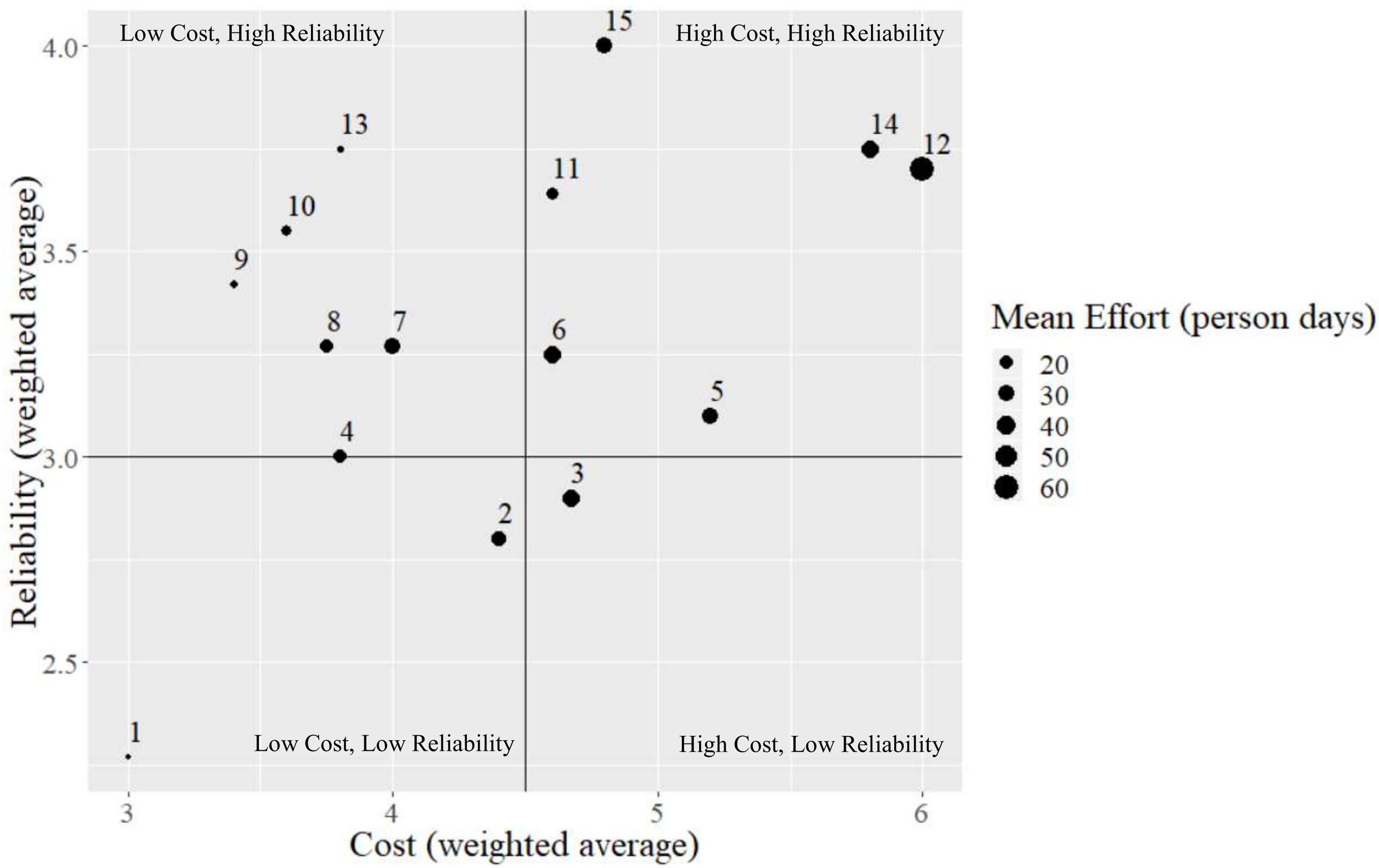


Figure 2.5. An influence diagram links the most important stressors to bald eagles in Southwest Alaska National Parks to feasible monitoring metrics and fundamental objectives, defined by the expert panel. Stressors are connected to feasible monitoring metrics, based on panelist opinion of responsiveness. Metrics are connected to fundamental objectives based on cost, effort, reliability, and sensitivity.

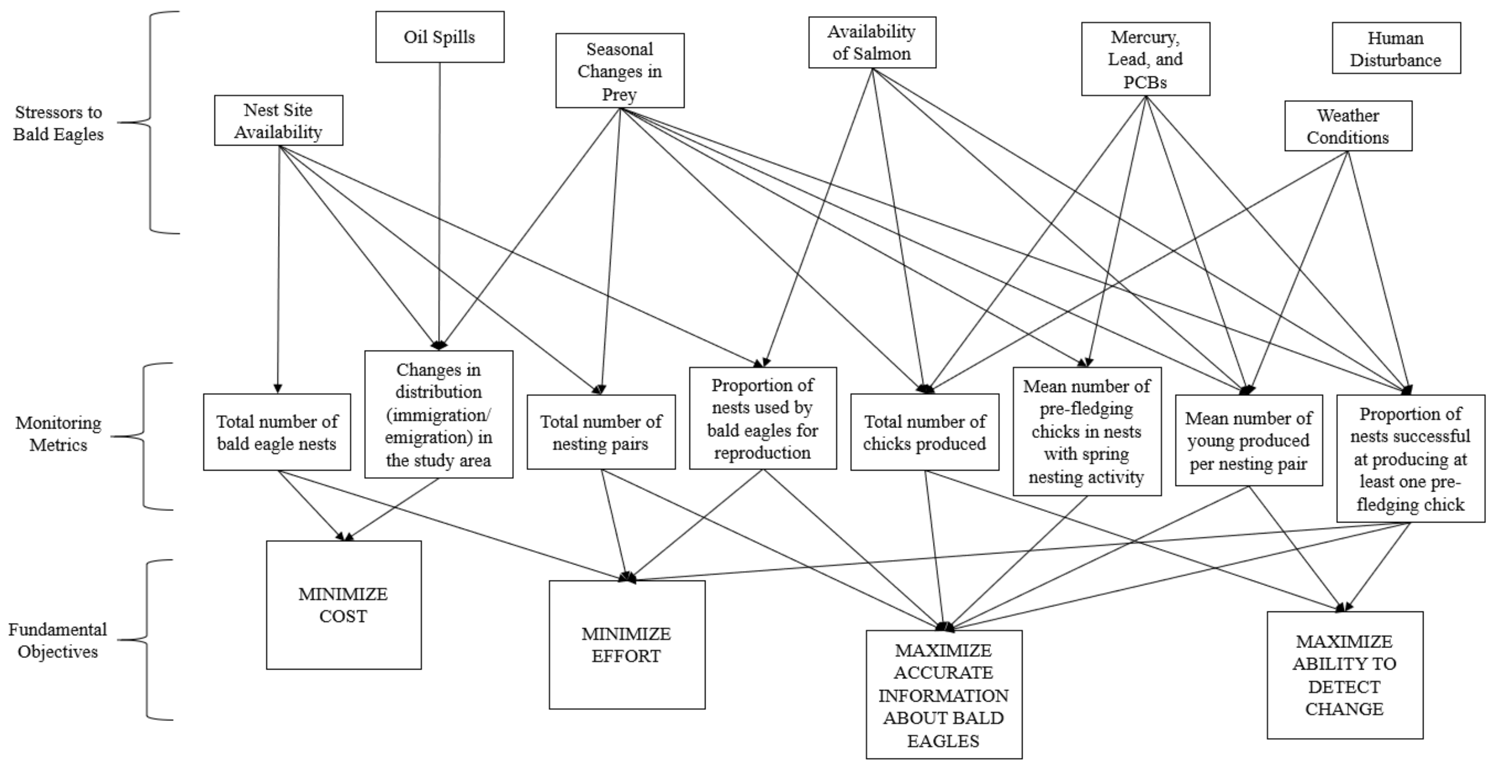


Table 2.1. Response numbers for each round of questioning are shown in this table. The percent response for each questionnaire is calculated using the number of responses and confirmed non-respondents. Percent Response is calculated using responses and confirmed non-responses.

Questionnaire	**Q1**	**Q2**	**Q3**	**Q4**
Responses	11	14	13	10
Responses + Confirmed Non-Respondents	12	15	14	12
Percent Response	66.67%	83.33%	77.78%	66.67%

Table 2.2. This table shows monitoring metric names, reliability scores, cost scores, and mean effort (estimated using annual person days). Reliability and Cost are given scores based on weighted averages from panelist responses. Effort is shown using mean annual person days, as estimated by panelists. Number of respondents is denoted in each cell by n.

Label	Metric	Reliability	Cost	Effort
1	Total number of Bald Eagle nests	2.27 (n=11)	3 (n=5)	11.5 (n=4)
2	Changes in distribution (immigration/emigration) in the study area	2.8 (n=10)	4.4 (n=5)	26.67 (n=3)
3	Abundance at winter concentration sites	2.9 (n=10)	4.67 (n=3)	33.33 (n=3)
4	Total number of chicks produced	3 (n=12)	3.8 (n=5)	23 (n=4)
5	Proportion of breeding to non-breeding eagles	3.1 (n=10)	5.2 (n=5)	28 (n=3)
6	Nestling survival from hatching to fledging	3.25 (n=12)	4.6 (n=5)	31.33 (n=3)
7	Mean number of young produced per nesting pair	3.27 (n=12)	4 (n=5)	30 (n=4)
8	Mean number of pre-fledging chicks produced in nests with spring nesting activity	3.27 (n=12)	3.75 (n=4)	20 (n=3)
9	Proportion of nests used by Bald Eagles for reproduction	3.42 (n=12)	3.4 (n=5)	13 (n=4)
10	Proportion of nests that are successful at producing at least one pre-fledging chick	3.55 (n=12)	3.6 (n=5)	16 (n=4)
11	Mean number of chicks fledged per nesting pair	3.64 (n=12)	4.6 (n=5)	19.5 (n=4)
12	Abundance of bald eagles as determined by statewide aerial surveys	3.7 (n=10)	6 (n=4)	63.33 (n=3)
13	Total number of nesting pairs	3.75 (n=12)	3.8 (n=5)	12 (n=4)
14	Subadult survival	3.75 (n=12)	5.8 (n=5)	35 (n=4)
15	Adult survival	4 (n=12)	4.8 (n=5)	27.5 (n=4)

Table 2.3. A simplified contingency table relates important stressors and feasible metrics, based on sensitivity of metrics, and the number of metrics that measure each stressor. The sum of each column shows sensitivity (how many stressors each metric is sensitive to) and the sum of each row shows the number of metrics that may measure each stressor.

Metrics / **Stressors**	Total number of Bald Eagle nests	Changes in distribution (immigration/ emigration) in the study area	Total number of nesting pairs	Proportion of nests used by Bald Eagles for Reproduction	Total number of chicks produced	Mean number of pre-fledging chicks produced in nests with spring nesting activity	Mean number of young produced per nesting pair	Proportion of nests that are successful at producing at least one pre-fledging chick	**Number of Metrics that Measure Each Stressor**
Availability of salmon	0	0	0	1	1	0	1	1	**4**
Nest site availability/suitable nesting locations	1	1	1	1	0	0	0	0	**4**
Seasonal changes in prey	0	1	1	0	1	1	1	1	**6**
Oil spills	0	1	0	0	0	0	0	0	**1**
Mercury, lead, and PCBs	0	0	0	0	1	1	1	1	**4**
Weather conditions	0	0	0	0	1	0	1	1	**3**
Human Disturbance	0	0	0	0	0	0	0	0	**0**
Sensitivity	**1**	**3**	**2**	**2**	**4**	**2**	**4**	**4**	

Table 2.4. A consequence table identifies which metrics should remain in consideration for the bald eagle monitoring program. Each metric is assigned a score for the four fundamental objectives, based on panelist responses. Metrics that are outcompeted by other metrics in all four categories are considered “Dominated” and are thus eliminated from consideration.

		A	B	C	D	E	F	G	H	I	J	K	L
Objective	**Performance Measure**	**Total # bald eagle nests**	**Changes in distribution**	**Abundance at winter concentration sites**	**Productivity**	**Proportion of breeding to non-breeding eagles**	**Nestling survival from hatching to fledging**	**Proportion of nests used by bald eagles for reproduction**	**Proportion of nests successful at producing at least one pre-fledging chick**	**Abundance of bald eagles as determined by statewide aerial surveys**	**Total # nesting pairs**	**Subadult survival**	**Adult survival**
Minimize Cost (↓)	Weighted average of annual cost	3	4.4	4.67	3.5	5.2	4.6	3.4	3.6	6	3.8	5.8	4.8
Minimize Effort (↓)	Mean # annual person days	11.5	26.67	33.33	15	28	31.33	13	16	63.33	12	35	27.5
Maximize Ability to Detect Change (↑)	Sensitivity	5	9	0	5.5	0	1	3	5	5	2	5	6
Maximize Accurate Information about Bald Eagles (↑)	Weighted average of reliability	2.27	2.8	2.9	3.7	3.1	3.25	3.42	3.55	3.7	3.75	3.75	4
Dominated by:				D,F,G,H,J		D,G,H,J,L	D,G,H,J		D	D,K,L		L	
Eliminate?		N	N	Y	N	Y	Y	N	Y	Y	N	Y	N

CHAPTER 3

USING A BAYESIAN NETWORK MODEL TO EVALUATE BALD EAGLE VITAL SIGNS MONITORING IN SWAN PARKS

ABSTRACT

Structured decision making approaches can be used in complex and contentious natural resource decisions that involve many stakeholders. By providing organized methods, structured decision making allows groups of decision-makers to reevaluate objectives and options, as well as analyze tradeoffs of various alternatives. The Southwest Alaska Network (SWAN) monitors bald eagles as part of the Vital Signs Monitoring Plan. Although bald eagles are abundant and appear to have relatively stable populations in this region, it is still important to monitor the species because of its inherent importance to park visitors and because of the bald eagle's role as an ecological indicator. We gathered an expert panel of scientists and managers, and implemented a Delphi Process to gather information about the bald eagle monitoring program. We also had panelists generate a list of fundamental objectives for the monitoring program: minimizing cost, minimizing effort, maximizing the ability to detect change in bald eagle populations, and maximizing the amount of accurate information collected about bald eagles. We used a swing weighting technique to assign weights for each fundamental objective. Collecting accurate information about bald eagles was considered the most important fundamental objective, while cost was considered the least important. Using the information collected in the Delphi Process along with the fundamental objectives and their corresponding weights, we analyzed the scenarios and defined the most optimal decision using a Bayesian Decision Net. Through our analysis, we found that a

"Comprehensive" monitoring scenario, comprised of all feasible monitoring metrics is the most optimal monitoring scenario. We found that even with a greatly increased cost, the Comprehensive monitoring scenario remains most optimal, however by increasing both cost and effort required for this scenario to 4.4 times the amount of cost and effort required for the current "Status Quo" program, the current monitoring scenario becomes the most optimal. We recommend further exploration of the exact cost and effort required for the Comprehensive scenario, to determine if it is in the parks' best interest to begin monitoring additional metrics.

INTRODUCTION

Decisions in the field of natural resources are fundamentally complex, due to, not only biological elements, but the social component that is intrinsically linked with fish and wildlife issues (Duda et al. 1998, Mendoza and Martins 2006). Structured decision making can help to conquer some of this complexity by encouraging decision makers to consider innovative solutions in a systematic and transparent manner. It focuses decision-makers on key questions about the context of the decision, objectives, uncertainties, and trade-offs of various alternatives, while highlighting the importance of continuously learning (Gregory et al. 2012).

Structured decision making is defined by Gregory et al. (2012) as "the collaborative and facilitated application of multiple objective decision making and group deliberation methods to environmental management and public policy problems." It can be compared to and fit into an adaptive framework, as both exhibit the similarities of defining explicit objectives and alternatives. Structured decision making approaches can serve as decision aids to facilitate monitoring programs that explicitly address the decisions being made, and can help to conserve limited resources by reducing the waste of time and effort (Lyons et al. 2008, Gregory et al. 2012). Ultimately, monitoring programs that spend an adequate amount of time defining objectives and optimizing the program are more successful, since their monitoring is focused on important data needs for conservation and wildlife issues (Oakley et al. 2003, Nichols and Williams 2006).

Ideally, structured decision making should be enacted at the conception of a monitoring program. Following the "roadmap" by Reynolds et al. (2016) for designing and implementing a monitoring program, an adequate program should include steps to encompass the general phases of framing the problem, designing the monitoring program, implementing and learning, and learning and revising. Steps to frame the problem and analyze alternatives should also be revisited frequently, since values and attitudes can change over the course of an extended or repeated decision (Williams 2011). Constantly revisiting decisions allows a monitoring program to be useful by collecting consistently relevant information that will relate directly to decisions that are being made (Nichols and Williams 2006) and to remain relevant with changing agency employees (Oakley et al. 2003). Instead, many programs begin by collecting data before laying the groundwork, and the value of the monitoring effort can be diminished (Reynolds et al. 2016). Gregory et al. (2012) state that "[a] casual approach to monitoring and to adaptive management, albeit widespread, is both naïve and wasteful." A structured approach to decision making facilitates transparent and comprehensive decisions regarding multi-objective problems (Martin et al. 2009, Gregory et al. 2012) and ultimately leads to a more efficient program by identifying the optimal survey design for monitoring (Reynolds et al. 2011). It also allows decision-makers to learn about and improve the decision process for continued improvement in management decisions (Williams 2011).

Decisions in the natural resource field, including those relating to monitoring, also often involve multiple stakeholders and decision-makers. These issues can be difficult to navigate and can further complicate the decision-making process. Unfortunately, collaborative decisions tend to be hindered by logistical constraints, which prevent improvements in monitoring (Reynolds et al. 2016). While it may be easier to shy away from these types of group decisions, hearing the opinions of multiple experts can encourage deeper thinking from individuals (Runge et al. 2011). It has also been demonstrated that airing and resolving conflict during the decision process (rather than after the decision is made) promotes satisfaction among the group about the final decision (Priem and Harrison 1995). Group decisions should be inclusive to a wide array of stakeholders, with a stakeholder defined as a person or group of people who has a vested interest in a particular issue or decision (Grimble and Wellard 1997). It is recommended that an open discourse be created and upheld between field scientists, managers, those analyzing the data, and other stakeholders throughout the decision making process to maintain support for decisions (Reynolds et al. 2011). By considering the wide range of values and interests, decision-making problems can be more rigorously explored and objectives can be made clearer. Stakeholder analysis is regarded as most valuable in situations where objectives and values among stakeholders are contentious (Grimble and Wellard 1997).

These complex decisions in natural resource management often couple group decision making with multiple-objective issues (Williams 2011, Gregory et al. 2012).

Objectives may be based on economic constraints, social ideals, and the value of collecting scientific information (Grimble and Wellard 1997). These objectives, along with potential alternatives, are formed and judged based on the values of the group of stakeholders responsible for making the decision, and may be disputed between stakeholder groups, especially in situations where resources and actions are limited (Nichols and Williams 2006, Lyons et al. 2008). A way to optimize the decision, or perform the actions that best meet all objectives collectively, will allow the data to be put to its best use for the purposes of conservation (Nichols and Williams 2006, Lyons et al. 2008). Realistically, multiple-objective decisions will involve tradeoffs or sacrifices for valuing one objective over another. Tradeoffs are inevitable each time a decision is made, but using stakeholder values in a structured decision making approach allows these tradeoffs to be highlighted to examine the cost of choosing one alternative over another (Grimble and Wellard 1997).

National Parks in Southwest Alaska face complicated decisions regarding monitoring of vital signs, which provide information about the health and stability of park ecosystems and resources (Bennett et al. 2006). Lake Clark National Park and Preserve, Katmai National Park and Preserve, Kenai Fjords National Park, and Wrangell – St. Elias National Park monitor bald eagles annually. During an in-person meeting of a panel of experts, the parks have defined fundamental programmatic objectives for a monitoring decision. Through the use of a Delphi Process, a panel of experts has analyzed a suite of monitoring metrics, and these metrics will be considered in their reevaluated monitoring program. Now, the decision of which monitoring metrics to use in the program must be optimized to consider the parks' objectives of minimizing cost, minimizing effort, maximizing the ability to detect change, and maximizing the amount of accurate information about bald eagles.

Bayesian Belief Network (BBN) models have been used to combine causal influence diagrams, scientific information, and expert opinion to predict the outcome of management actions or decisions. These models can use several different types of data and can easily be altered and updated. A variation of the Bayesian Belief Network is a Bayesian Decision Net, which aids in decision-making by determining the utility of various alternatives, based on objectives (Marcot et al. 2001). Provided there is a clear plan for how the Bayesian Decision Net will be used, this method facilitates an unambiguous way to make difficult decisions that balance multiple objectives and highlights the consequences and trade-offs of each option (Marcot et al. 2006, Fortin et al. 2016).

By balancing fundamental objectives in a decision model, each alternative will be assigned a utility and the best decision can be made, given the current information and the values and needs of all stakeholders. These alternatives were formed through a structured survey technique, the Delphi Process, that uses surveys to combine expert opinion. Now that monitoring metrics have been evaluated by experts, we will use the information we collected about each monitoring metric and current monitoring information to optimize the decision. Stakeholders have already developed fundamental objectives for the monitoring program and performance measures with which to measure success in achieving them. Based on these objectives of minimizing cost, minimizing effort, maximizing accurate information collected about bald eagles, and maximizing the ability to detect change in bald eagle populations, a decision will be made to select a set of monitoring metrics that balances these needs. This method to select an optimal bald eagle monitoring program is an example of using structured decision making techniques to formally and transparently analyze complex problems and make a decision that combines the opinions of many experts.

METHODS

We gathered an expert panel of 18 scientists, managers, and personnel from the National Park Service, US Fish and Wildlife Service, and South Dakota Game Fish & Parks to participate in a Delphi Process. We queried the panel about long-term bald eagle monitoring in Southwest Alaska National Parks, and about the cost, effort, reliability, and sensitivity of various monitoring metrics. Through an in-person panel meeting, we formed fundamental objectives for bald eagle monitoring program decisions in SWAN parks: Minimize Cost, Minimize Effort, Maximize Ability to Detect Changes in Bald Eagle Populations, Maximize Accurate Information about Bald Eagles. We evaluated monitoring metrics, based on those fundamental objectives, using inputs from panelists and a consequence table to eliminate metrics that were outcompeted by other metrics. A full description of these methods can be found in Chapter 2. The monitoring metrics that remain in consideration after this process are: total number of bald eagle nests, changes in distribution, productivity, proportion of nests used by bald eagles for reproduction, total number of nesting pairs, and adult survival.

Although there are a wide variety of alternative scenarios that can be formed using subsets of the remaining six metrics, we chose six scenarios to represent feasible options for monitoring. The scenario "Status Quo" includes feasible metrics that are currently monitored by the parks. The "Comprehensive" scenario consists of all six metrics designated feasible by the expert panel. There is also a scenario, “No Monitoring” that considers the option to discontinue monitoring bald eagles. “New Metrics” considers metrics that are feasible, but not currently monitored by the parks (adult survival and changes in distribution). There are also two scenarios “Reduced Status Quo 1” and “Reduced Status Quo 2” that consider currently monitored metrics with a reduced monitoring effort. Table 3.1 outlines the metrics that are included in each scenario. We designed these scenarios to cover a range of reasonable options that are comprised of the feasible metrics identified by the expert panel.

For each fundamental objective for the bald eagle monitoring program (minimize cost, minimize effort, maximize accurate information about bald eagles, maximize ability to detect changes in bald eagle populations), we assigned values for each scenario to evaluate the performance of that scenario with respect to each objective. We generated these values using information collected from the expert panel during the Delphi Process and using budget values from SWAN parks.

Cost is assigned using panelist response from the Delphi Process. We asked panelists to assign a cost value to each metric, for each annual year of surveying. We gave multiple choice options for each metric: $0-5,000; $5,000-10,000; $10,000-15,000; $15,000-20,000; $20,000-25,000; $25,000+. Panelist responses were combined into a weighted average value, with higher values representing higher cost for that metric. We assigned an effort score to each scenario using responses from the Delphi Process. We asked experts to estimate annual person days required for each individual metric, and calculated the mean for each metric. For each scenario, we added mean annual effort values for individual metrics that comprise the scenario. The amount of accurate information about bald eagles is measured based on a reliability rating, gathered during the Delphi Process. A reliability score was assigned to each metric, through the use of multiple choice questions administered to panelists. We asked panelists to rate the reliability of metrics on a 5-point scale and calculated the weighted average for each metric; this weighted average represents the reliability score. We added these values for the metrics that comprise each scenario, to create a reliability score for each scenario. The ability to detect change is measured using a sensitivity score. To create this sensitivity score, we used information collected during the Delphi Process: experts were asked to select metrics that are responsive to important stressors to bald eagles. The sensitivity score for each metric is a count of the stressors to which that metric is responsive. For each scenario, we added the sensitivity scores of the metrics that comprise the scenario.

For the two reduced effort Status Quo scenarios, we did not collect information about the fundamental objectives directly from the expert panel. Aided by an expert, we assigned values to these scenarios based on their relative performance to the Status Quo scenario. Since the Status Quo scenario is comprised of three annual surveys, we estimated that one third of each score is attributed with each annual survey. Using these approximations, we calculated scores for scenarios by eliminating part of one annual survey and eliminating an entire annual survey. We also supplemented estimations with current park budgets. We normalized values for each scenario, and those normalized values are used in the decision model. Table 3.2 displays the value assigned to each scenario for the four fundamental objectives.

We determined the weight of fundamental objectives, based on importance. These weights are determined by the panel of experts using a swing-weighting technique, adapted from the USFWS/USGS Structured Decision Making Workshops course (Course code: ALC3159). The form distributed to the expert panel is shown in Table 3.3 (this table includes responses from just one panelist as an example). We distributed a personalized form to each panelist using Google Sheets. We listed each fundamental objective along with corresponding performance metrics, and whether our aim is to maximize or minimize that attribute. We displayed a range of values, including the worst and best possible values for each attribute. We based the worst and best possible values on panelist responses to the Delphi Process questionnaires. We also displayed five hypothetical scenarios. A “Benchmark” scenario is comprised of the worst possible values for all four fundamental objective attributes. In the remaining four hypothetical scenarios, all attributes are set to their worst values except for one attribute in each scenario, which is set to its best value.

We asked panelists to rank the four hypothetical scenarios from 1-4 (1 is best). The Benchmark scenario is automatically assigned the worst rank of 5. By doing this, we are asking the panelists which attribute they would swing to its best level, if they could only pick one. That scenario receives the rank of 1. The next most important swing is ranked 2, etc. We then asked panelists to score each scenario based on its priority. The Rank 1 scenario automatically receives a score of 100. Panelists assigned scores in decreasing amounts to the remaining hypothetical scenarios based on importance in achieving each measure swing. We provided the example to panelists that if they score their Rank 2 scenario at 50, they are saying that it is half as important to achieve that measure swing as the measure swing in their Rank 1 scenario, which has a score of 100. Using Equation 1, we assigned a weight to each fundamental objective for each individual panelist and created box and whisker plots for each objective. To combine panelist responses for cumulative objective weights that will be used in the decision model, we averaged individual panelist *weight (normalized)* values for each fundamental objective.

Equation 1. *This equation is used to calculate normalized weights for fundamental objectives.*

$$weight\ (normalized) = \left[\frac{score}{sum\ of\ scores}\right] * 100$$

We then combined these elements to create our decision model, using program Netica from Norsys Software Corp. to create a Bayesian Decision Net. The decision net uses three types of nodes: a decision node, nature nodes, and a utility node. The decision node allows the user to select a scenario alternative and displays the utility value of each scenario. The decision node connects to the nature nodes. We created a nature node for each fundamental objective. These nature nodes are thus named “Cost”, “Effort”, “Accurate_Info”, and “Detect_Change”. Using the normalized score values for each objective, we populated the model in Netica. These values are then routed through the utility node, which uses the equation shown below, in Equation 2.

Equation 2. *This equation for a linear value model calculates a utility value for each alternative monitoring scenario.*

$$Utility = \sum W_i X_i$$

$$where\ W_i = the\ weight\ of\ fundamental\ objective\ i, and$$

$$X_i = the\ performance\ score\ for\ each\ fundamental\ objective$$

This equation uses a technique called Linear Value Modeling, as described by Gregory et al. (2012). The value model incorporates the weight assigned to each objective by expert panelists, using the swing weighting technique. These weights may be changed to examine the effect that changing values may have on the decision outcome. Our linear value model used (*1 – normalized value*) for “Cost” and “Effort” since our goal is to minimize these attributes. We used the normalized values for “Accurate_Info” and “Detect_Change” since our goal is to maximize these attributes. The utility values are displayed in the decision node. The scenario with the highest utility value is considered the most optimal decision.

We also performed a sensitivity analysis to examine differences in individual stakeholder values, using methods from Converse et al. (2013). They outline a sensitivity analysis that determines scenario rankings for each panelist’s responses, individually. They determine if any individual’s outcome differs from the outcome using consensus objective weights, which are averaged across panelists. We ran the decision model using each individual panelists’ assigned weights.

We performed an additional sensitivity analysis to determine the change in objective weights needed to alter the outcome of the decision model. We added utility values for all scenarios to create a total utility. We calculated the percentage of each scenario utility compared to the total utility. We also calculated the percent utility of each scenario, disregarding one fundamental objective at a time. We presented these results graphically and visually examine them to determine sensitivity. To identify the point at which scenario rankings change, we examined where lines intersect on the graphs. By setting linear models for various scenarios equal and solving for the objective weight, we determined at which weight one scenario begins to outcompete another.

We also examined the sensitivity of objective weights as if this were a two-factor decision model, instead of a four-factor decision model. We combined the objective weights and utility values for cost and effort, since their assigned weights were very similar. We then examined sensitivity by graphing the percent total utility for each scenario across various objective weights, with cost and effort combined. We then combined objective weights and utility values for Accurate Information and Detect Change, since their assigned weights were highly correlated. We examined the sensitivity by graphing the percent total utility for each scenario across various objective weights, with Accurate Info and Detect Change combined. By setting linear equations for various scenarios equal and solving for objective weight, we determined at which weight one scenario began to outcompete another.

Due to concerns about underestimating the costs of measuring adult survival, we tested additional scenarios with an increased cost for measuring adult survival and changes in distribution. Using Millsap et al. (2002) as a reference for appropriate transmitters and Bowman et al. (1995) and Buehler et al. (1991) as reference data for the number of eagles that would be tagged, we calculated a rough estimate of the cost for monitoring adult survival in bald eagles. Microwave Telemetry avian transmitters, which are suitable to monitor bald eagles, range in base price from $2900 - $3650 per transmitter. Bowman et al. (1995) tagged 79 adult bald eagles in Prince William Sound, Alaska to study survival following the *Exxon Valdez* oil spill. Additional bald eagle survival studies tagged 70 eagles (Millsap et al. 2002) and 39 eagles (Buehler et al. 1991).

We used the minimum and maximum transmitter prices as a proxy for cost. To tag 39 eagles, transmitters alone would cost between $113,100-$142,000. Using these same cost values, transmitter costs to tag 79 eagles would cost between $229,100-$288,350. Assuming there would be additional unaccounted costs to monitor adult survival, we estimated price values for the New Metrics scenario, and adjusted the price value for the Comprehensive monitoring scenario. We estimated a new price value for the New Metrics scenario, assuming 39 eagles are tagged, as $200,000. We estimated a new price value for the New Metrics scenario, assuming 79 eagles are tagged, as $300,000. We adjusted price values accordingly for the Comprehensive monitoring scenarios. We then recalculated normalized cost values and ran the decision model using these values. We ran the Bayesian Decision Net using the newly calculated normalized cost values for tagging 39 and 79 bald eagles. We graphically examined the sensitivity of each fundamental objective weight, using the increased cost values.

Since the Comprehensive monitoring scenario is the only scenario that outcompetes the Status Quo monitoring scenario in the decision model, we graphically examined scenarios that may cause the Comprehensive scenario to be outcompeted. We explored the role of increased cost of the Comprehensive scenario in terms of proportional cost value (*Cost of Comprehensive Scenario/Cost of Status Quo Scenario*). We aimed to determine how much larger the proportional cost must be so that the Status Quo scenario outcompetes the Comprehensive scenario, given all other performance measures remain the same. We created a graph that compares the proportional value of scenario costs on the x-axis (*Cost of Comprehensive Scenario/Cost of Status Quo Scenario*) and the proportional utility value on the y-axis (*Utility of Comprehensive Scenario/Utility of Status Quo Scenario*). By doing so, we compared how much higher the comprehensive scenario must be in cost for it to fall below the status quo scenario in ranking, given all other fundamental objective measures and objective weights remain the same. Where this line crosses the threshold of a proportional value of 1, the Status Quo scenario begins to outcompete the Comprehensive scenario.

We also examined the role of increasing cost and effort of the Comprehensive monitoring scenario to determine how much larger the Comprehensive scenario must be in cost and effort than the Status Quo scenario for it to be outcompeted. We tested various increased cost values for the Comprehensive scenarios, to the point where the cost of the Comprehensive scenario is 500 times larger than the Status Quo scenario. We also increased effort in a proportional manner to the increased cost. For example, if the increased cost value for the Comprehensive scenario was three times larger than the Status Quo scenario, we multiplied the Status Quo effort value by three to get the increased effort value. We calculated the proportional value of the Comprehensive scenario cost plus effort values to the Status Quo cost and effort values. We also calculated the proportional Utility value of the Comprehensive monitoring scenario to the Status Quo monitoring scenario. Where this line crosses a horizontal threshold of 1, the Status Quo monitoring scenario begins to outcompete the Comprehensive monitoring scenario.

RESULTS

Abbreviated names of scenarios can be found in Table 3.1. Values assigned to each fundamental objective for all scenarios are displayed in Table 3.2. We used normalized scores in the decision model. For the objectives “Minimize Cost” and “Minimize Effort”, we used the values *1-normalized score* in the decision model, since the aim was to minimize these attributes.

Figure 3.1 is a radar chart that displays the rank of each scenario for each fundamental objective, separately. For each fundamental objective, the scenarios that score higher are considered better in that category than scenarios that score lower. For the fundamental objective “Minimize Cost” (displayed as “Cost”), the No Monitoring scenario scores the highest since it costs nothing. For the Cost objective, following No Monitoring in decreasing order: Reduced SQ2, New Metrics, Reduced SQ1, Status Quo, and Comprehensive. For the fundamental objective “Minimize Effort” (displayed as “Effort”), the No Monitoring scenario scores the highest since it requires no effort. For the Effort objective, following No Monitoring in decreasing order: Reduced SQ2, Reduced SQ1, Status Quo, New Metrics, and finally Comprehensive. For the objective “Maximize Ability to Detect Change” (displayed as “Detect Change”), the Comprehensive scenario scores the highest. For the Detect Change objective, following Comprehensive in decreasing order: Status Quo, New Metrics, Reduced SQ1, Reduced SQ2, and finally No Monitoring. For the objective “Maximize Accurate Information about Bald Eagles” (displayed as “Accurate Info”), the Comprehensive scenario scores the highest. For the Accurate Info objective, following Comprehensive in decreasing order: Status Quo, Reduced SQ1, Reduced SQ2, New Metrics, and finally No Monitoring.

Panelists assigned individual weights to each of the four fundamental objectives, based on their perception of its importance to the monitoring program. Box and whisker plots of panelist objective weights are found in Figure 3.2. Cost weights ranged from 8.3% to 30.2%. Effort weights ranged from 8% to 28.3%. Detect Change weights ranged from 18.9% to 41.7%. Accurate Info weights ranged from 26.7% to 40%. Mean weights, calculated from individual panelist weights, were used in the decision model. Accurate Info had the highest mean weight (33.1%), followed by Detect Change (32.3%), Effort (17.6%), and Cost (17.1%).

The Bayesian Decision Net calculates the overall utility score for each scenario by combining objective scores and fundamental objective weights. The Bayesian Decision Net is displayed in Figure 3.3. Based on the scores assigned to each scenario for the four fundamental objectives, the Comprehensive monitoring scenario has the highest utility value, making it the most optimal monitoring scenario. The Comprehensive monitoring scenario has a utility score of 65.35. It is followed by the Status Quo scenario (54.30), Modified Status Quo 1 (51.31), Modified Status Quo 2 (48.30), New Metrics (45.71), and No Monitoring (34.55). Based on scores and objective weights, No Monitoring is the least optimal monitoring scenario.

For each scenario, the breakdown of utility scores by fundamental objective is displayed in Figure 3.4. The most optimal solution, according to this model, is comprised of scores from Detecting Change and collecting Accurate Information. Although it received scores of zero for cost and effort, it still outranked all other scenarios.

Results of the sensitivity analyses examining fundamental objective weights are displayed in the graphs in Figure 3.5. Objective weights are varied along the x-axis. On the y-axis, we compare the percentage of each scenario utility to the total utility of all scenarios added together. By visually inspecting these graphs, we can determine the sensitivity of each objective to changes in weight. Each time scenario lines intersect, the ranking of the most optimal scenarios changes. Where these lines intersect along the x-axis denotes the weight that the objective must reach for the ranking of scenarios to change. By increasing the weight of the cost and effort scenarios, there are changes in scenario rankings. For most objective weights, either the Comprehensive monitoring scenario or No Monitoring scenario is ranked highest. At very low objective weights for cost, the Comprehensive monitoring scenario performs the best of all scenarios. Once the value of cost increases to a weight of 34.4%, intermediate scenarios are most optimal, until cost is valued at a weight of 37.9%, when No Monitoring becomes most optimal. Similarly, if effort is valued at low objective weights, the Comprehensive monitoring scenario outcompetes all other scenarios, until it reaches an objective weight of 31.7%. At that point, the Status Quo scenario outcompetes all other scenarios, and falls below Reduced Status Quo 2 for a very narrow window, until No Monitoring begins to outcompete all other monitoring scenarios at an objective weight of 40.8%.

There are few changes in scenario rankings when varying Accurate Information and Detect Change objectives. These objective weights must decrease to low values for the Comprehensive scenario to be outcompeted by another monitoring scenario. All monitoring scenarios have similar utilities until Accurate Info reaches an objective weight of 4.7%, when the Comprehensive scenario quickly outcompetes all other scenarios. When varying the objective weight for Detect Change, the Status Quo scenario outcompetes all other metrics until the objective weight reaches 17.3%, when the Comprehensive scenario tops the rank of monitoring scenarios.

By combining objective weights and utility values for Cost and Effort, we observed similar sensitivity as for Cost or Effort individually (Figure 3.6). At low objective weights for Cost + Effort, the comprehensive scenario has a much higher % Utility than other scenarios. At an objective weight of 65.8%, the Comprehensive scenario is no longer the top competing model. At an objective weight value of 79.9%, the No Monitoring scenario becomes the highest ranked scenario for high objective weights of Cost + Effort.

In the same regard, we observed similar sensitivity for the combined objective weights of Accurate Info and Detect Change as we did for either objective individually (Figure 3.6). The Comprehensive scenario outcompetes all other scenarios for all objective weights of 24.4% and higher for Accurate Info + Detect Change. However, at lower objective weights than 24.4%, the Comprehensive scenario performs relatively poorly.

The values used in the increased cost models are shown in Table 3.4. By increasing the cost of the new metrics to $200,000, the Comprehensive scenario remains the highest ranked monitoring scenario, followed by the Status Quo scenario. Increasing the cost of the new metrics to $300,000 does not change the scenario ranking, and the Comprehensive scenario is still the highest ranked monitoring scenario, followed by Status Quo. The results of the Bayesian Decision Nets using increased cost scenarios are shown in Figures 3.7 and 3.8. By comparing the proportional utility value of the comprehensive scenario to the status quo scenario and the proportional cost of these two scenarios, we find that even when the cost of the Comprehensive monitoring scenario is more than 500 times larger than the Status Quo scenario, the Comprehensive monitoring scenario still outcompetes the Status Quo scenario, provided all other utility values and all objective weights remain the same for the Comprehensive scenario (Figure 3.9).

By evaluating the effects of increased cost and effort, we find that the combined cost and effort value must be 4.4 times larger for the Comprehensive monitoring scenario than the Status Quo scenario, for the Status Quo scenario to outcompete the Comprehensive model (Figure 3.10).

DISCUSSION

Structured decision making helps to add rigor and reflection into scientific monitoring programs. We evaluated several monitoring scenarios for bald eagles in four national parks in Alaska, based on fundamental objectives defined by relevant experts. Although it requires the maximum amount of cost and effort of any scenario option, our decision model identifies the most optimal monitoring scenario as a comprehensive program. This includes monitoring for all feasible metrics to gather the most information about bald eagles as possible and maximize the ability to detect changes in the population.

While any number of monitoring scenarios could have been analyzed using these methods, we chose six scenarios that we felt adequately represent the range of options that vary in the cost and effort they require, as well as the information they provide. Following guidelines from Gregory et al. (2012), the alternatives that should be considered when making a decision should be able to provide a complete and meaningful resolution to the problem at hand. Therefore, we did not consider alternatives that would clearly be unreasonable. A Status Quo scenario includes all the feasible metrics that are currently monitored, while a Comprehensive scenario includes all six metrics deemed feasible in previous analysis. Since the two new metrics that were suggested are high in cost and effort, we included them together as a unique scenario, “New Metrics”. Although it would provide no information about the bald eagle population, “No Monitoring” is a decision that could be made, so it was included as a scenario. We explored a reduced status quo scenario, which removed half of the second May occupancy survey. A second reduced status quo scenario completely removed the second May occupancy survey. These scenarios would reduce the cost and effort of the monitoring effort, while still providing an adequate amount of information and ability to detect change, although reduced from the current monitoring program. We calculated the cost and effort for these scenarios based on a previous budget, and the values entered for amount of accurate information and ability to detect change were estimated with the help of an expert. While we feel these values adequately represent these reduced scenarios, if parks choose to pursue these options, more rigorous statistical methods should be used to obtain more accurate values for these scenarios.

Competing fundamental objectives make it necessary to consider tradeoffs, which are inevitable in natural resource decisions (Gregory et al. 2012, Converse et al. 2013). Our decision model helped to quantify those tradeoffs, and identified the comprehensive monitoring scenario as the most optimal model. Given the weight of the fundamental objectives, the increased ability of this scenario to collect accurate information about bald eagles and detect changes in the population greatly outweighed the negative tradeoffs of high cost and effort. Although the status quo scenario scored relatively well and outcompeted most other scenarios, the addition of adult survival to the comprehensive model caused it to outweigh the status quo scenario. Both reduced status quo scenarios were eliminated due to the reduction in accurate information and ability to detect change, despite a reduced cost and effort, which had relatively low importance to the expert panelists. Although monitoring just the newly suggested metrics, adult survival and changes in distribution, would provide a significant amount of information and ability to detect change, the cost and effort are not low enough for this scenario to outweigh the current monitoring regime. No monitoring was eliminated since it would provide no information about bald eagles and provide no ability to detect changes in the populations. This would leave the parks in a difficult situation if a problem arose and bald eagle populations began to decline.

The results of our decision model relied heavily on the weights assigned by panelists. Since cost and effort were both valued at relatively low weights, the utility of each scenario was largely based on its ability to collect accurate information about bald eagles and detect changes in populations. The monitoring scenario that ultimately had the largest overall utility received scores of 0 for cost and effort, but was high enough in its ability to collect accurate information and detect change that it outcompeted all other monitoring scenarios. Just as Gende et al. (2018) explore the balance between and resources and park values in management of human visitors, this decision balances the resources needed for a monitoring program and the values of park scientists: the information collected from the bald eagle monitoring program. In the case of our decision model, the balance between resources and information leans heavily toward the data collected on bald eagles, thus determining the outcome of the decision model.

It became apparent while examining the fundamental objective weights that this can be viewed as a two-function decision model, rather than a four-function decision model. Although panelists specified four fundamental objectives of minimizing cost, minimizing effort, maximizing the amount of accurate information collected, and maximizing the ability to detect changes in the population, there is a clear distinction in the objective weights. Cost and effort were weighted very similarly, as were detecting change and collecting accurate information. Perhaps this decision is realistically based on two categories, broadly: resources required, and information obtained. We confirmed this conjecture by examining the sensitivity of objective weights in a two-factor decision model, with cost and effort combined into a single objective, and accurate info and detect change combined into a second objective. As this discussion of an optimal monitoring program continues, it may be beneficial to reduce the decision to a more simplistic cost-benefit analysis, with cost and effort combined into a "resources required" objective and accurate info and detect change combined into an "information obtained" objective. While simplifying the problem to two fundamental objectives would still be considered a multi-criteria decision analysis, the decision can be improved with greater simplicity (Mendoza and Martins 2006).

For sensitivity analyses that vary cost and effort objective weights (either separately or as a combined objective), there are only very narrow ranges of weights that allow intermediate scenarios to outcompete the Comprehensive or No Monitoring scenarios. Given our input values, by varying objective weights of cost and effort, this decision about monitoring is largely "all or nothing," with either the Comprehensive or No Monitoring scenarios as the highest ranked options for all objective weights outside of these narrow ranges. For our decision model, there would need to be a huge shift in values (either much more value assigned to cost and effort, or much less value assigned to accurate information and detecting change) for scenarios to begin outcompeting the Comprehensive scenario.

The scores assigned to each objective and used as inputs in the decision models were generated using expert opinion, including that of decision makers. This is a valid method of collecting information to supplement empirical data (MacMillan and Marshall 2006, Eycott et al. 2011), help to identify expected value and uncertainty of information (Runge et al. 2011), and to evaluate the values involved in a monitoring decision (Gregory et al. 2012). However, it also introduces a substantial amount of uncertainty. Especially for the monitoring metric "Adult Survival," there is quite a bit of ambiguity in all four performance measures. Since survival is not currently monitored, we asked panelists to make educated guesses about the values that should be entered in the model. Since methods to measure adult survival were not specified, panelists were likely considering differing methods when estimating performance measures. As an example, one panelist suggested measuring survival by collecting feathers from tree bases, and thus had lower estimates of cost and effort. Additionally, the way the question about cost was presented limited the cost of monitoring Adult Survival to a value that was likely much less than the realistic cost. Finally, we had low response numbers to questions asking panelists to estimate cost and effort of various monitoring metrics. Some panelists felt unqualified to make those estimates, and these questions were asked late in the Delphi

Process when response rates tended to be lower. Since, for some measures of cost and effort, we had as few as three respondents estimating values of performance measures, these values may not be as accurate as other estimates. Estimates of cost and effort varied greatly.

Since estimates for Adult Survival cost varied so widely, we thought it was important to analyze the effects of increased cost for this metric. Although the cost to monitor adult survival in bald eagles in such a remote and rugged landscape would likely be higher than our estimate, we chose a conservative estimation of cost for this metric when examining the consequences. Even with substantial cost increases of greater than 10 times the original panelist estimate, the Comprehensive monitoring scenario continues to have the highest utility. The Status Quo scenario, which was the next highest performing scenario under for all of our model runs, can have a cost that is much smaller than the Comprehensive monitoring scenario and still be outcompeted. However, when effort also increases along with increasing cost, these two factors together must be 4.4 times larger for the Comprehensive scenario than the Status Quo scenario, for the Comprehensive scenario to lose its designation of most optimal model. This is an important consideration, since cost and effort of monitoring adult survival and changes in distribution could easily amount to greater than 4.4 times larger than the current monitoring program. It is also worth noting that any change in monitoring program will require additional resources in the form of cost and effort to write new protocols, train staff, create new data templates, and analyze and interpret data.

We speculate that the weights assigned to fundamental objectives are heavily influenced by the fact that all panelists completing the ranking and scoring are biologists.

Biologists, as opposed to higher level management, are most likely biased to care more about the persistence of a species and value information about a biological system more than the cost and effort it requires. It would provide an interesting perspective to present the swing-weighting form to a panel of directors and budgetary decision makers to see how the weights of objectives changes, and we recommend that this be explored before making any changes to the bald eagle monitoring program.

In conclusion, this process displayed the inherent importance of gathering accurate information and being able to detect changes in the populations of symbolic and charismatic wildlife populations. Even in relatively undisturbed wildlife populations, such as the bald eagle populations in Southwest Alaska National Parks, park scientists are driven by the need to maintain an adequate bank of information about the species and its health. As an important ecological indicator, the bald eagle can provide knowledge about the health of the park ecosystems. Although there is a desire to reduce the resources required to monitor bald eagles in the parks: Katmai National Park and Preserve, Kenai Fjords National Park, Lake Clark National Park and Preserve, and Wrangell – St. Elias National Park and Preserve, the structured decision making process highlighted the relative unimportance of cost and effort in determining the most optimal monitoring program. Although our decision model repeatedly identified a Comprehensive monitoring scenario, consisting of all feasible monitoring metrics, as most optimal, we recommend a more thorough analysis into the exact needs of the parks to monitor adult survival and changes in distribution, and to generate a more specific cost estimate to analyze the tradeoffs involved in taking on this more intensive monitoring effort.

LITERATURE CITED

Bennett, A. J., W. L. Thompson, and D. C. Mortenson. 2006. Vital Signs Monitoring Plan, Southwest Alaska Network. National Park Service, Anchorage, AK.

Bowman, T. D., P. F. Schempf, and J. A. Bernatowicz. 1995. Bald Eagle Survival and Population Dynamics in Alaska after the Exxon Valdez Oil Spill. The Journal of Wildlife Management 59:317-324.

Buehler, D. A., J. D. Fraser, J. K. Seegar, and G. D. Therres. 1991. Survival Rates and Population Dynamics of Bald Eagles on Chesapeake Bay. The Journal of Wildlife Management 55:608-613.

Converse, S. J., C. T. Moore, M. J. Folk, and M. C. Runge. 2013. A matter of tradeoffs: Reintroduction as a multiple objective decision. The Journal of Wildlife Management 77:1145-1156.

Duda, M. D., S. J. Bissell, and K. C. Young. 1998. Wildlife and the American Mind. Responsive Management, Harrisonburg, VA.

Eycott, A. E., M. Marzano, and K. Watts. 2011. Filling evidence gaps with expert opinion: The use of Delphi analysis in least-cost modelling of functional connectivity. Landscape and Urban Planning 103:400-409.

Fortin, J. K., K. D. Rode, G. V. Hilderbrand, J. Wilder, S. Farley, C. Jorgensen, and B. G. Marcot. 2016. Impacts of Human Recreation on Brown Bears (Ursus arctos): A Review and New Management Tool. PLoS One 11:e0141983.

Gende, S., A. N. Hendrix, and J. Schmidt. 2018. Somewhere between acceptable and sustainable: When do impacts to resources become too large in protected areas? Biological Conservation 223:138-146.

Gregory, R., L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson. 2012. Structured Decision Making: A Practical Guide to Environmental Management Choices. John Wiley & Sons, West Sussex, UK.

Grimble, R., and K. Wellard. 1997. Stakeholder Methodologies in Natural Resource Management: a Review of Principles, Contexts, Experiences and Opportunities. Agricultural Systems 55:173-193.

Lyons, J. E., M. C. Runge, H. P. Laskowski, and W. L. Kendall. 2008. Monitoring in the Context of Structured Decision-Making and Adaptive Management. Journal of Wildlife Management 72:1683-1692.

MacMillan, D. C., and K. Marshall. 2006. The Delphi process - an expert-based approach to ecological modelling in data-poor environments. Animal Conservation 9:11-19.

Marcot, B. G., R. S. Holthausen, M. G. Raphael, M. M. Rowland, and M. J. Wisdom. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. Forest Ecology and Management 153:29-42.

Marcot, B. G., J. D. Steventon, G. D. Sutherland, and R. K. McCann. 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Canadian Journal of Forest Research 36:3063-3074.

Martin, J., M. C. Runge, J. D. Nichols, B. C. Lubow, and W. L. Kendall. 2009. Structured decision making as a conceptual framework to identify thresholds for conservation and management. Ecological Applications 19:1079-1090.

Mendoza, G. A., and H. Martins. 2006. Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms. Forest Ecology and Management 230:1-22.

Millsap, B. A., T. Breen, E. McConnell, T. Steffer, L. Phillips, N. Douglass, and S. Taylor. 2002. Comparative Fecundity and Survival of Bald Eagles Fledged from Suburban and Rural Natal Areas.pdf. Florida Fish and Wildlife Conservation Commission.

Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. Trends Ecol Evol 21:668-673.

Oakley, K. L., L. P. Thomas, and S. G. Fancy. 2003. Guidelines for Long-Term Monitoring Protocols. Wildlife Society Bulletin (1973-2006) 31:1000-1003.

Priem, R. L., and D. A. Harrison. 1995. Structured Conflict and Consensus Outcomes in Group Decision Making. Journal of Management 21:691-710.

Reynolds, J. H., M. G. Knutson, K. B. Newman, E. D. Silverman, and W. L. Thompson. 2016. A road map for designing and implementing a biological monitoring program. Environmental Monitoring and Assessment 188:399.

Reynolds, J. H., W. L. Thompson, and B. Russell. 2011. Planning for success: Identifying effective and efficient survey designs for monitoring. Biological Conservation 144:1278-1284.

Runge, M. C., S. J. Converse, and J. E. Lyons. 2011. Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. Biological Conservation 144:1214-1223.

Williams, B. K. 2011. Adaptive management of natural resources--framework and issues. J Environ Manage 92:1346-1353.

FIGURES AND TABLES

Figure 3.1. This radar chart displays normalized scores for each fundamental objective. For the cost and effort objective, this chart displays *1-normalized value* so that higher scores on this chart represent better-performing scenarios for each objective.

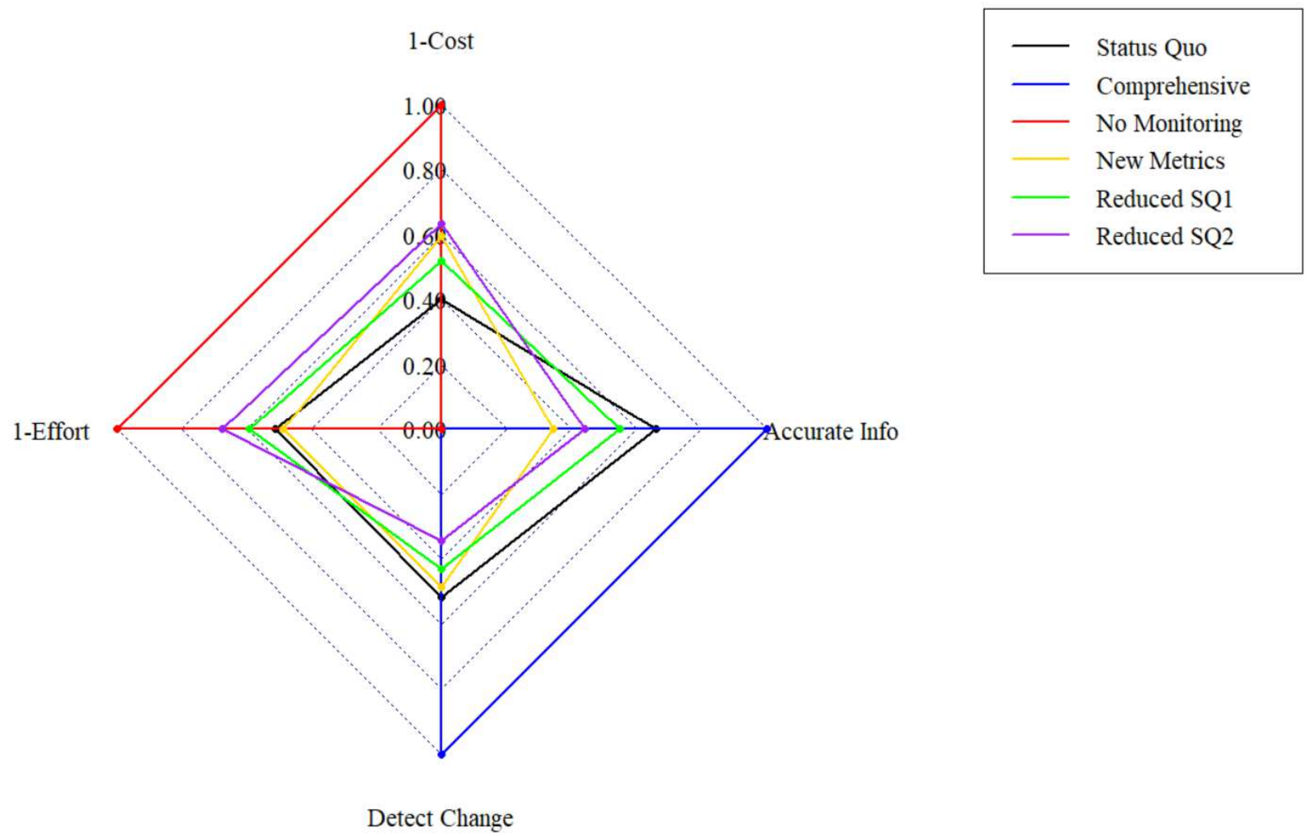


Figure 3.2. This series of boxplots shows the distribution of panelist weights (*n=10*) for the four fundamental objectives of the montioring program. These weights were collected through a swing-weighting procedure and average weights are used in the final decision model. A boxplot is shown for each objective, and the median is displayed on each plot. Outliers are represented by dots. The range of each boxplot represents the range of individual panelist responses.

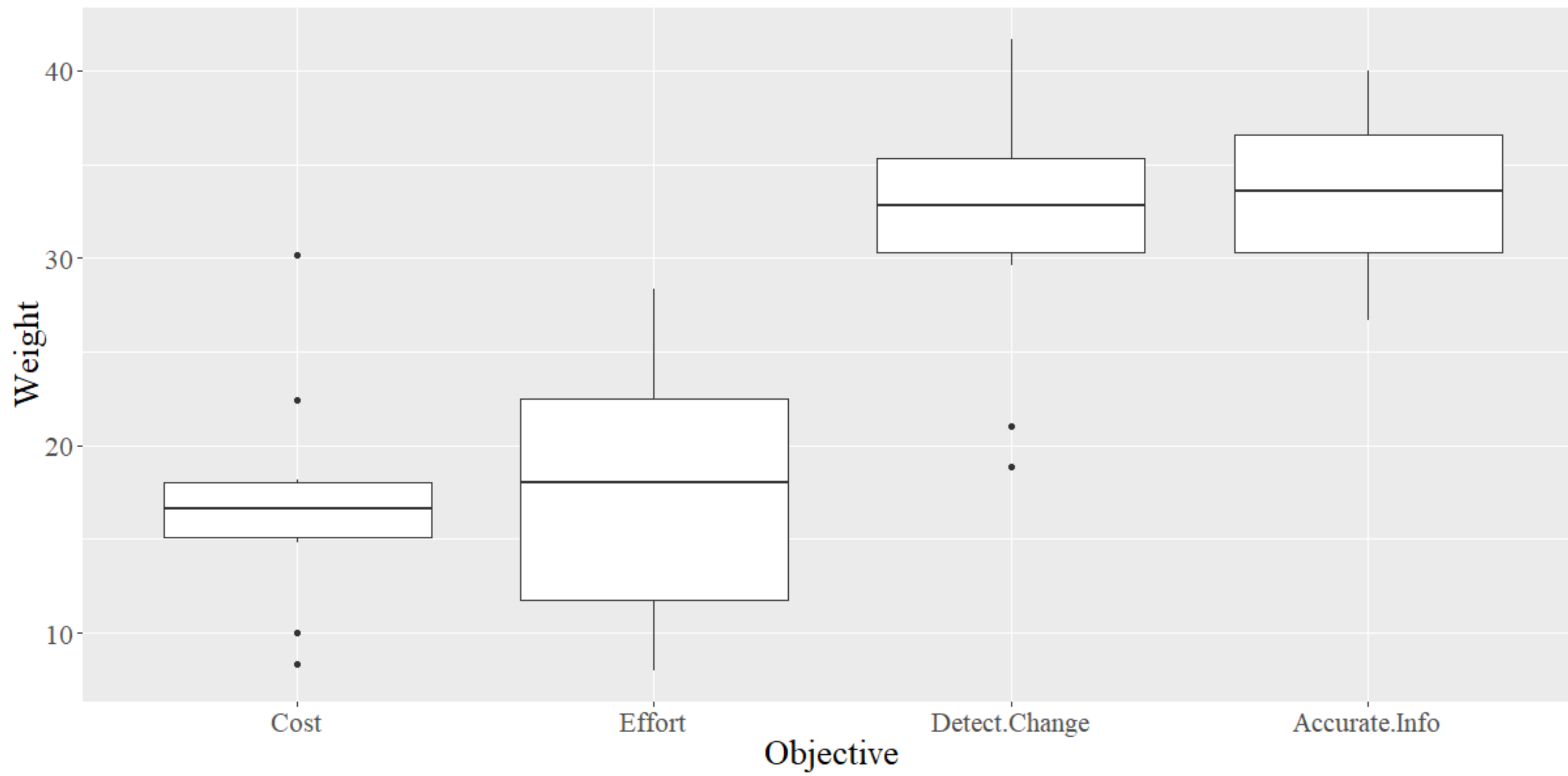


Figure 3.3. A Bayesian Decision Net uses a linear value modeling equation to combine scenario scores and fundamental objective weights into a utility score. Scenarios are displayed with their utility scored in the Scenario_Selection box. The boxes for each fundamental objective are compiled with scenario scores for each objective. The Utility box combines the data and provides the score in the Scenario_Selection box. The Comprehensive monitoring scenario has the highest utility value, making it the most optimal decision.

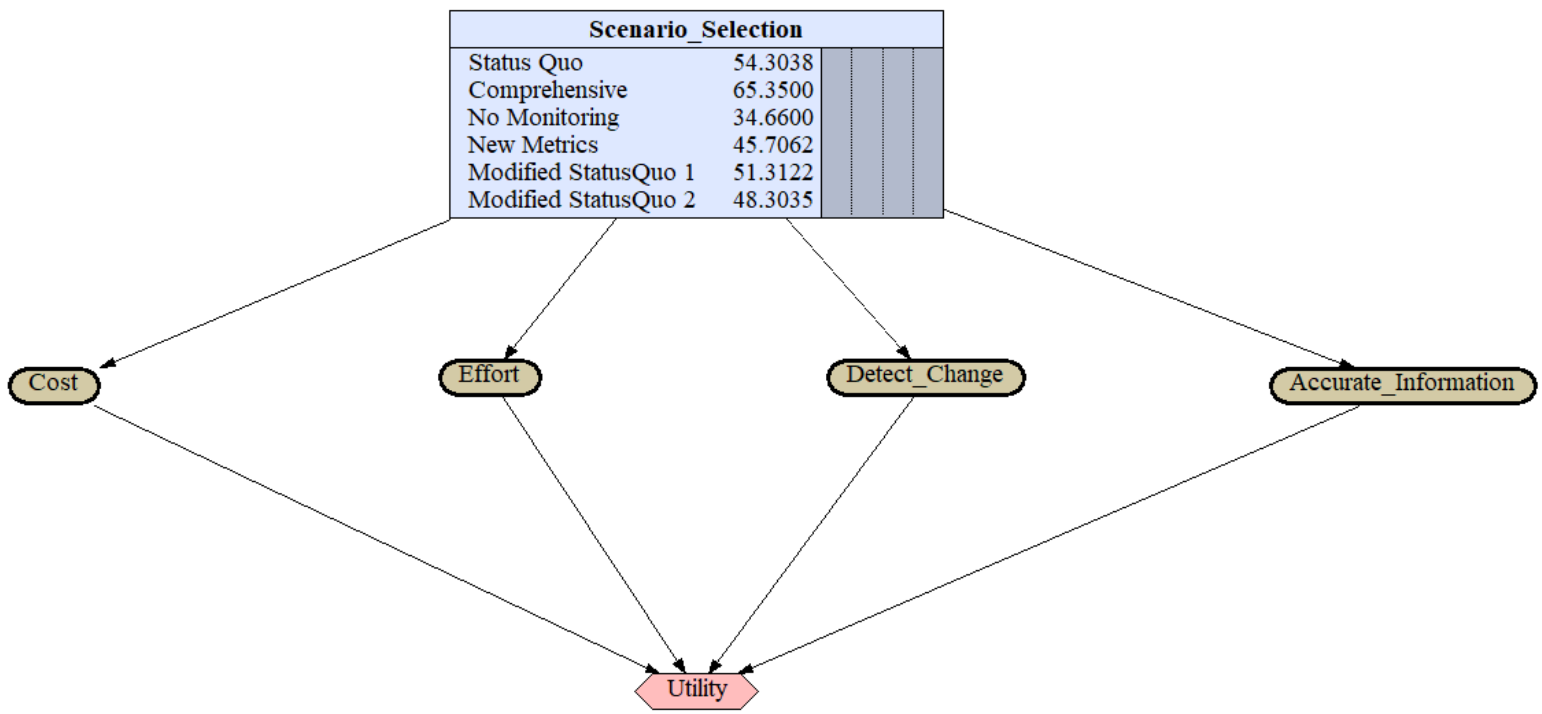


Figure 3.4. A bar chart shows the breakdown of overall utility scores by fundamental objective. The scores displayed in each portion of the bars are calculated by multiplying the scenario's score for that fundamental objective by the fundamental objective weight. Objective scores for each scenario add to that scenario's total utility score.

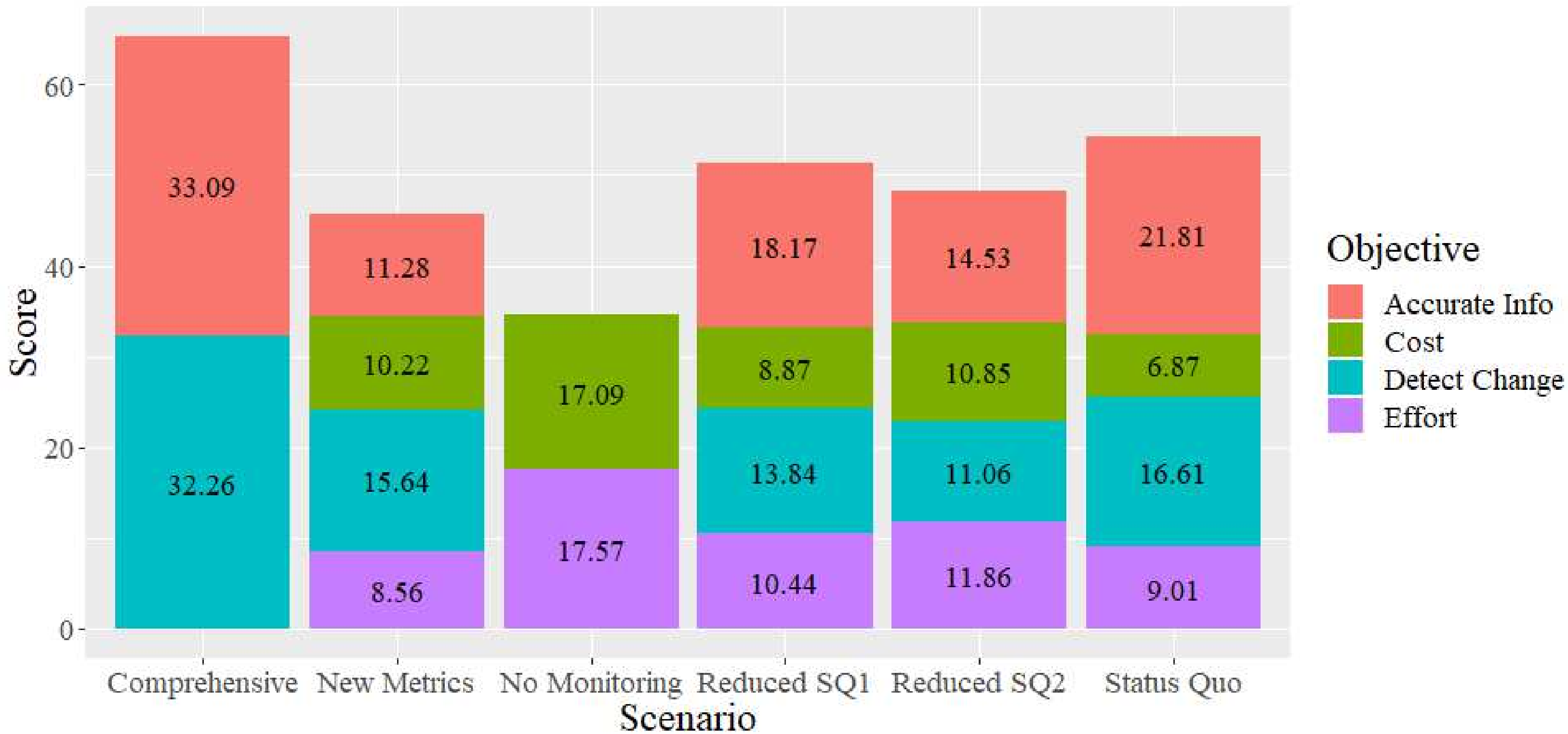
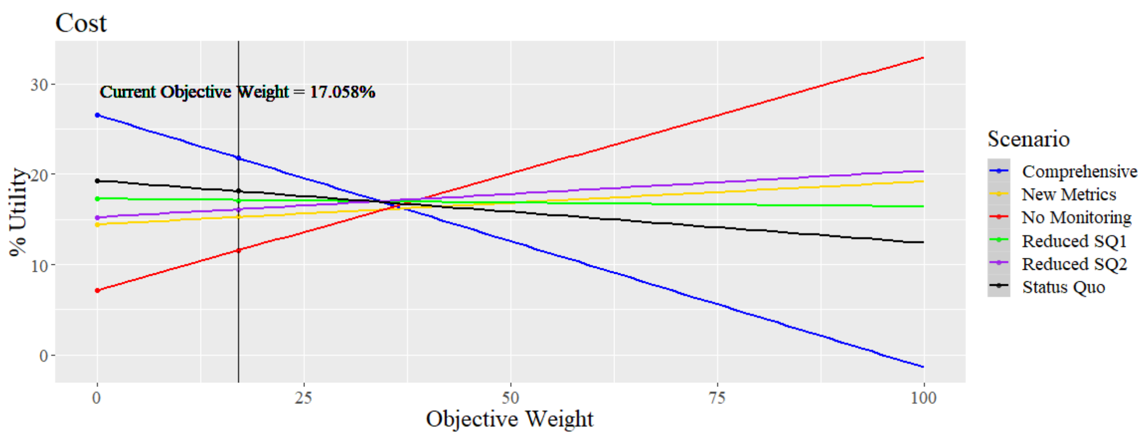
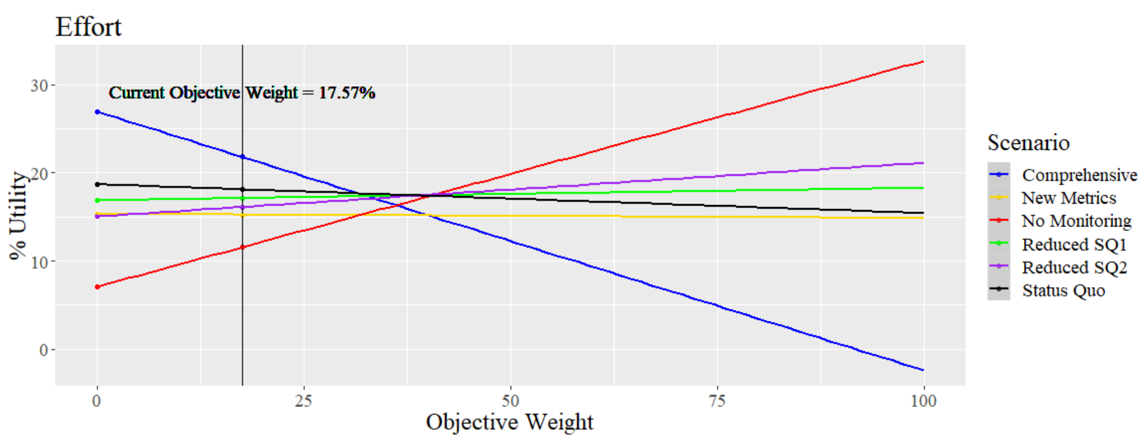
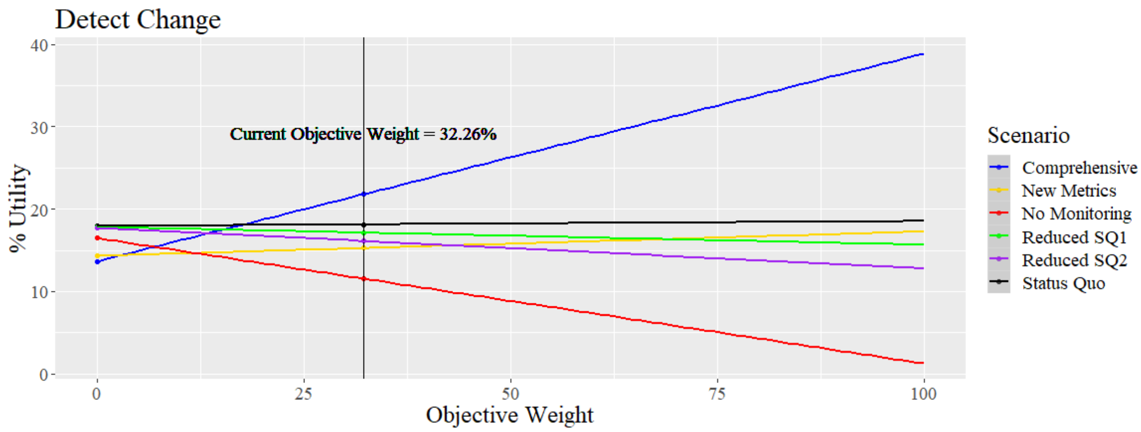


Figure 3.5. The sensitivity graphs show the response of % utility to changes in each objective weight. The % utility for each scenario is calculated by dividing the scenario's utility value at a particular objective weight by the total utility of all scenarios combined. Current objective weights, as defined by panelists, are labeled with a vertical line. The most optimal scenarios at a particular objective weight are designated by the order of the scenario lines, descending vertically. At current objective weights, the scenarios in order from most optimal to least optimal are: 1) Comprehensive; 2) Status Quo; 3) Reduced SQ1; 4) Reduced SQ2; 5) New Metrics; 6) No Monitoring.






Detect Change
Current Objective Weight = 32.26%
% Utility
Objective Weight
0
25
50
75
100
0
10
20
30
40
Scenario
Comprehensive
New Metrics
No Monitoring
Reduced SQ1
Reduced SQ2
Status Quo

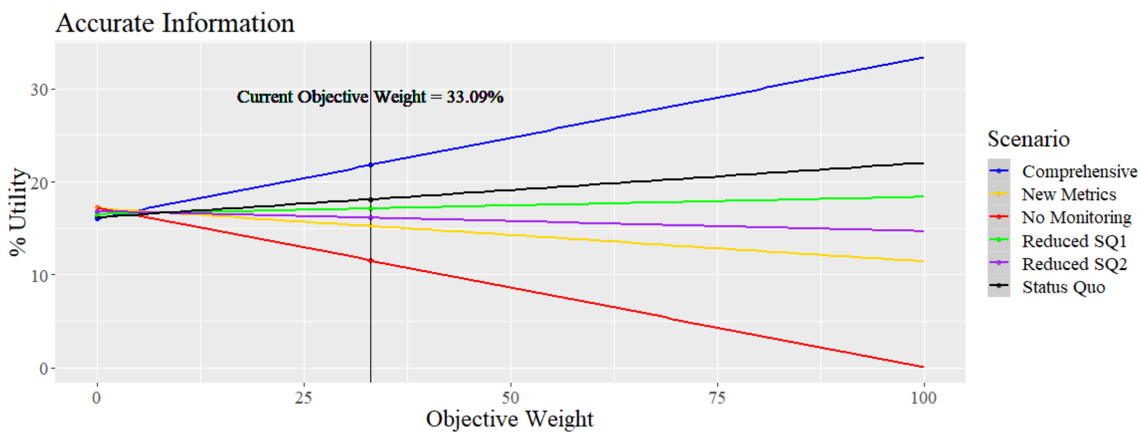

Accurate Information
Current Objective Weight = 33.09%
% Utility
Objective Weight
0
25
50
75
100
0
10
20
30
Scenario
Comprehensive
New Metrics
No Monitoring
Reduced SQ1
Reduced SQ2
Status Quo

Figure 3.6. The sensitivity graphs show the response of % utility to changes in each objective weight in a simplified manner, with Cost and Effort combined into one objective, and Accurate Info and Detect Change combined into another objective. The % utility for each scenario is calculated by dividing the scenario's utility value at a particular objective weight by the total utility of all scenarios combined. Current objective weights, as defined by panelists, are labeled with a vertical line. The current objective weight for Cost + Effort is the objective weights of cost and effort combined; the current objective weight for Accurate Info + Detect Change is the objective weights of these two individual objectives combined. The most optimal scenarios at a particular objective weight are designated by the order of the scenario lines, descending vertically. At current objective weights, the scenarios in order from most optimal to least optimal are: 1) Comprehensive; 2) Status Quo; 3) Reduced SQ1; 4) Reduced SQ2; 5) New Metrics; 6) No Monitoring.

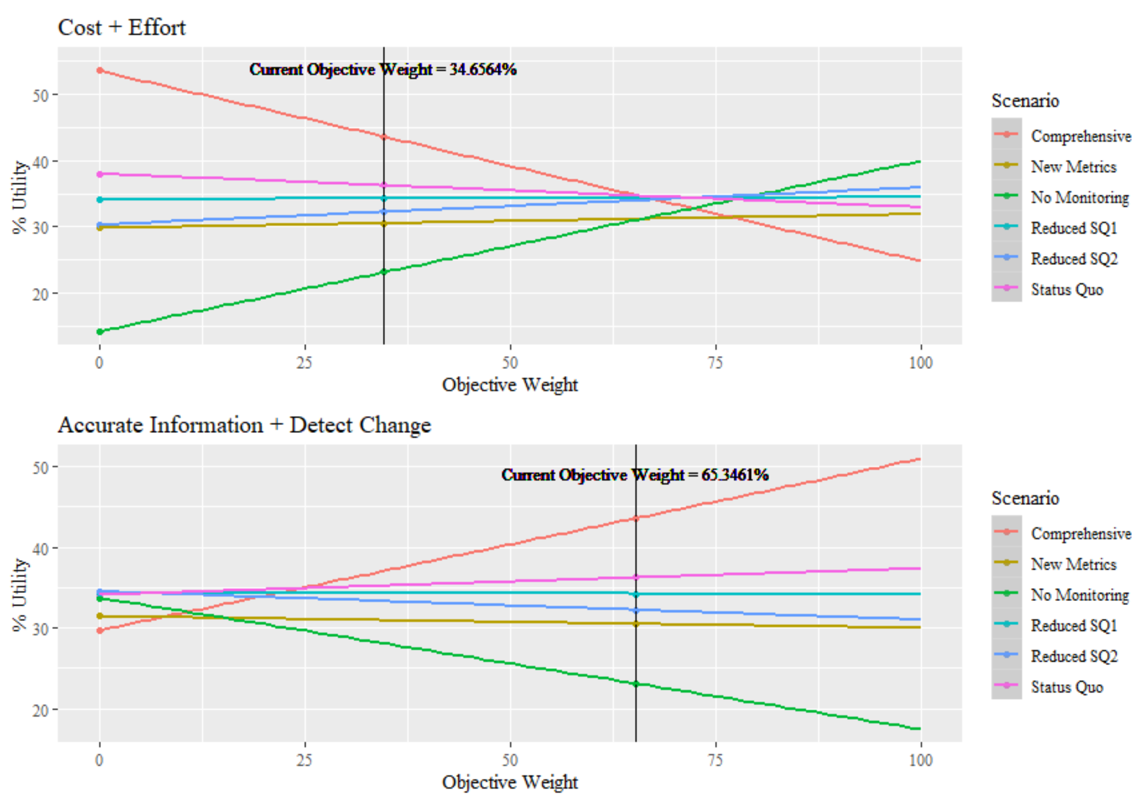


Figure 3.7. This decision model incorporates cost values associated with monitoring adult survival using 39 eagles. The values used in this decision model are displayed in Table 3.3.

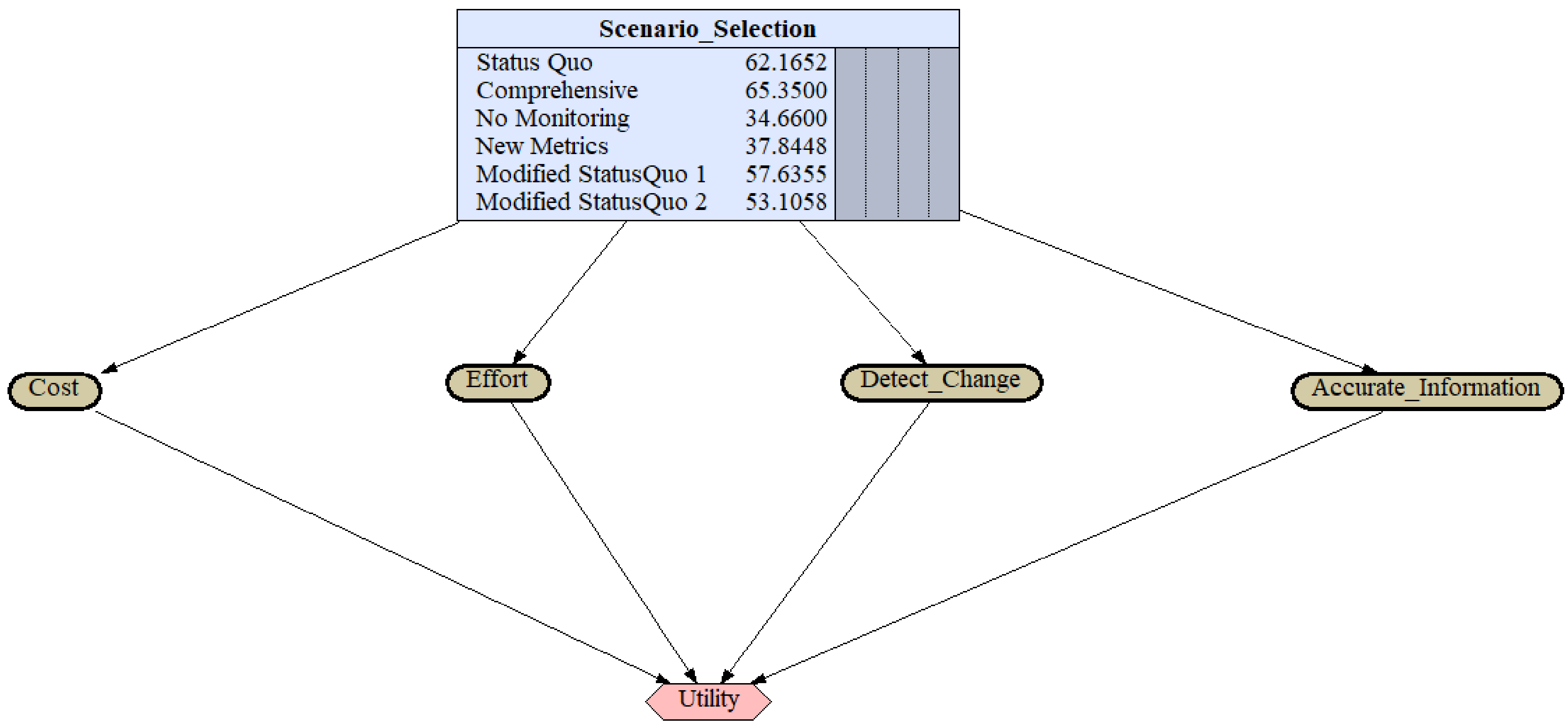


Figure 3.8. This decision model incorporates cost values associated with monitoring adult survival using 79 eagles. The values used in this decision model are displayed in Table 3.3.

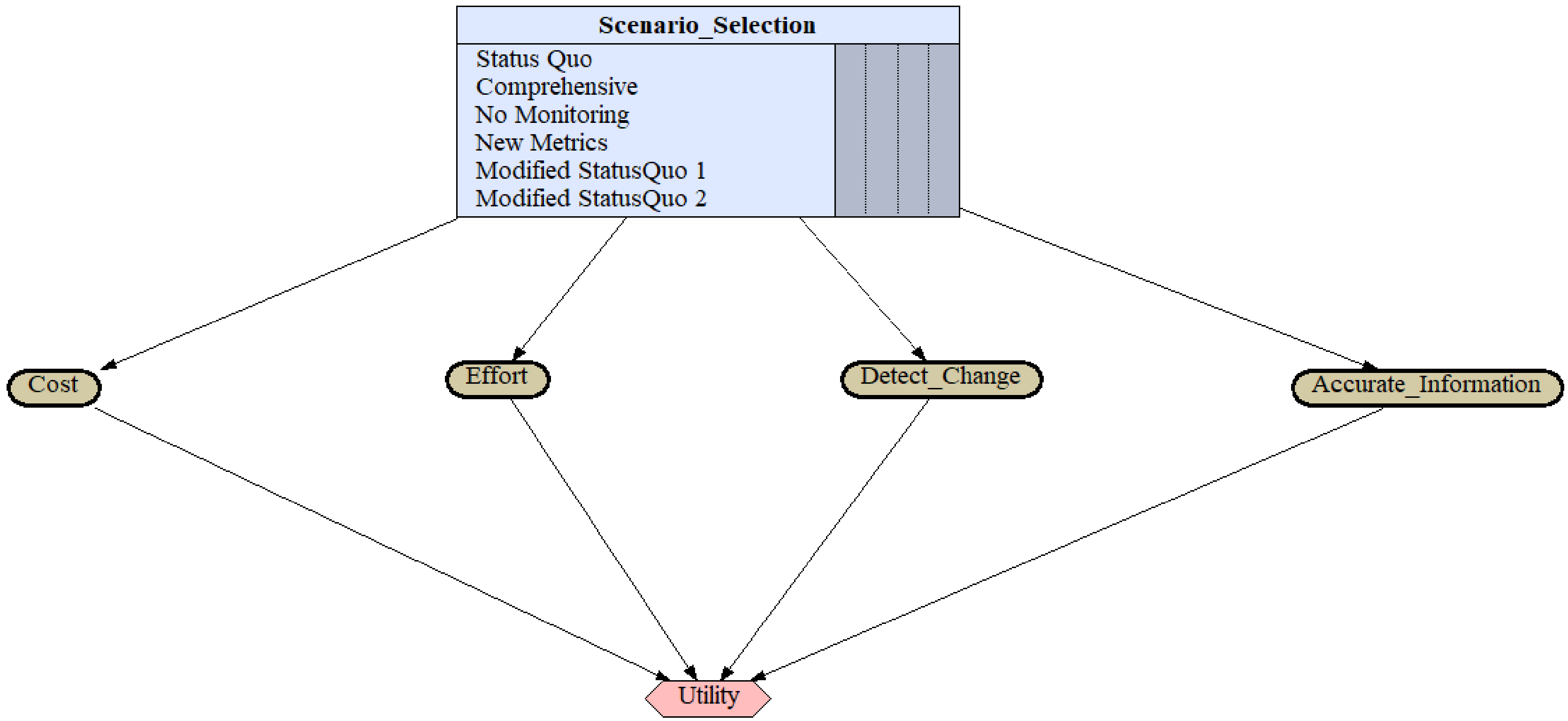


Figure 3.9. The comparison of proportional cost (Comprehensive Cost/Status Quo Cost) and proportional utility score (Comprehensive Utility/Status Quo Utility) is shown. When the curve crosses the horizontal line at Proportional Utility = 1, the Status Quo scenario would begin to outcompete the Comprehensive scenario. Even by increasing the cost of the Comprehensive monitoring scenario to 500 times the size of the Status Quo monitoring scenario, the Status Quo scenario still does not outcompete the Comprehensive monitoring scenario.

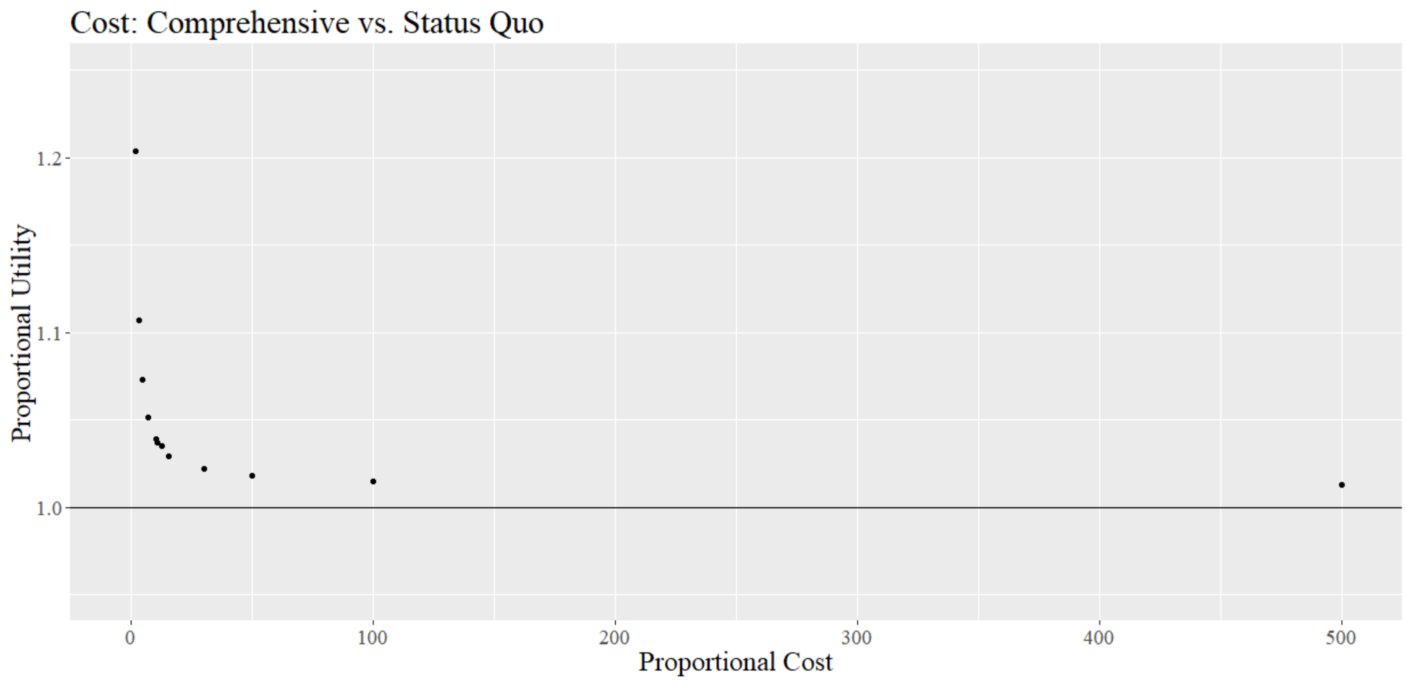


Figure 3.10. A curve representing increasing cost and effort values of the Comprehensive monitoring scenario to the Status Quo monitoring scenario. The Status Quo scenario begins to outcompete the Comprehensive scenario where the curve falls below the horizontal threshold of Proportional Utility = 1. At a point where cost and effort for the Comprehesive scenario are 4.4 times larger than for the Status Quo scenario, the Status Quo scenario begins to outcompete the Comprehensive scenario.

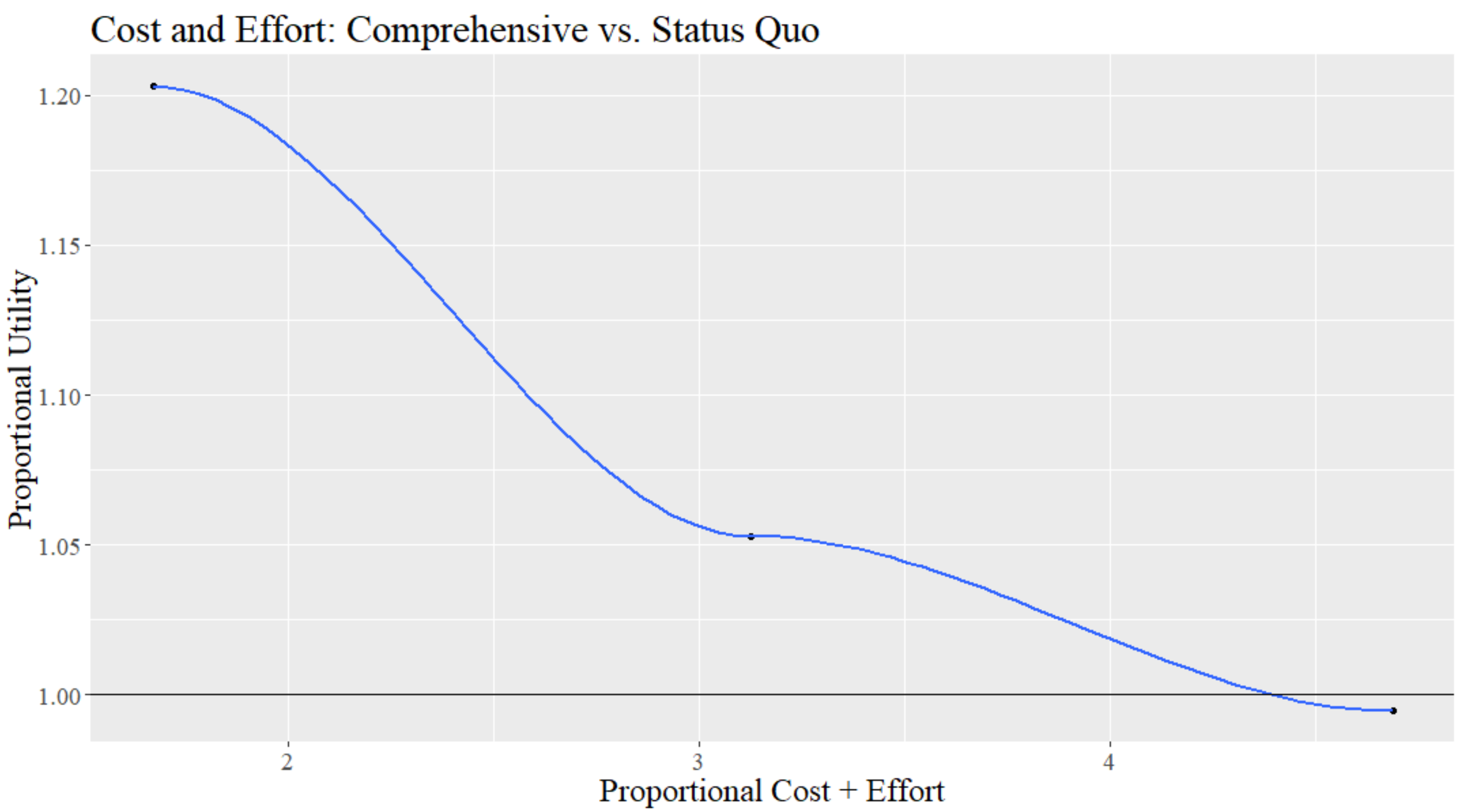


Table 3.1. This table displays the metrics that are included in each scenario in the decision model. These metrics were considered feasible by the expert panel. "Modified Status Quo 1" and "Modified Status Quo 2" incorporate the same metrics as "Status Quo" but remove various amounts of survey effort. Scenario Abbreviations may be used in subsequent figures and tables.

Scenario	Scenario Abbreviation	Metrics
Status Quo	N/A	Productivity Total number of bald eagle nests Proportion of nests used by bald eagles for reproduction Total number of nesting pairs
Comprehensive	N/A	Total number of bald eagle nests Changes in distribution Productivity Proportion of nests used by bald eagles for reproduction Total number of nesting pairs Adult survival
No Monitoring	N/A	N/A
New Metrics	N/A	Changes in distribution Adult survival
Modified Status Quo 1 (Remove 1/2 Second May Survey)	Reduced SQ1	Productivity Total number of bald eagle nests Proportion of nests used by bald eagles for reproduction Total number of nesting pairs
Modified Status Quo 2 (Remove Second May Survey)	Reduced SQ2	Productivity Total number of bald eagle nests Proportion of nests used by bald eagles for reproduction Total number of nesting pairs

Table 3.2. These values are incorporated into the decision model and signify the score each scenario is assigned for the four fundamental objectives. (1 – *normalized cost score*) and (1 – *normalized effort score*) are included because the linear value model uses these values, since these objectives are being minimized.

Scenario	**Cost Score**	**Normalized Cost Score**	**1 – (*normalized cost score*)**	**Estimated Budget Cost**	**Effort Score (days)**	**Normalized Effort Score**	**1 – (*normalized effort score*)**	**Detect Change Score**	**Normalized Detect Change Score**	**Accurate Info Score**	**Normalized Accurate Info Score**
Status Quo	13.7	0.598	0.402	$32,000	51.5	0.487	0.513	15.9	0.515	13.14	0.659
Comprehensive	22.9	1	0	$53,489*	105.57	1	0	30.9	1	19.94	1
No Monitoring	0	0	1	$0	0	0	1	0	0	0	0
New Metrics	9.2	0.402	0.598	$21,489*	54.17	0.513	0.487	15	0.485	6.8	0.341
Reduced Status Quo 1	11.02*	0.481	0.519	$25,750*	42.92*	0.406	0.594	13.25**	0.429	10.95**	0.549
Reduced Status Quo 2	8.35*	0.365	0.635	$19,500*	34.33*	0.325	0.675	10.6**	0.343	8.76**	0.439

*Calculated based on relative value

**Estimated based on relative value

Table 3.3. An example of a panelist's completed swing-weighting form to rank fundamental objectives by importance. Panelists completed the yellow-highlighted cells. In the hypothetical scenarios, all attributes for fundamental objectives were set to their worst-case value, except for one different attribute in each scenario. Panelists ranked the scenarios by most desired (1) to least desired (5). The Benchmark scenario was automatically given a rank of least desired (5). Panelists then scored scenarios. They were asked to automatically give their top-ranked scenario a score of 100, and assigned decreasing score values for the remaining scenarios. If they gave a scenario a score of 50, they are stating that they care about achieving that measure swing half as much as their top ranked scenario.

	Objective			**Range**		**Hypothetical Scenarios**				
	Description	**Attribute**	Goal	**Worst**	**Best**	**Benchmark**	**1**	**2**	**3**	**4**
A	Cost	Annual Dollars	min	$25,000+	$10,000	$25,000+	$10,000	$25,000+	$25,000+	$25,000+
B	Effort	Annual Person Days	min	63.33	11.5	63.33	63.33	11.5	63.33	63.33
C	Accurate Information about Bald Eagles	Reliability	max	2.27	4	2.27	2.27	2.27	4	2.27
D	Ability to Detect Change	Sensitivity	max	0	9	0	0	0	0	9
	Rank	1 is best; 5 (Benchmark) is worst				5	4	3	2	1
	Score	100 is best; 0 is worst. Assign 100 to Rank 1				0	25	50	75	100
	Weight (normalized)	[score/(sum of scores)]*100				0	10.000	20.000	30.000	40.000

Table 3.4. These values are assigned to increased cost scenarios, as a result of increasing the price of monitoring adult survival and changing distribution. Budget costs are estimated using the current bald eagle monitoring budget. Values that are estimated based on reference values use actual transmitter pricing data to approximate a cost. Normalized scores are entered in the decision model, but scenarios with higher values for *1 – (normalized objective score)* perform better, since cost is minimized in the linear value model.

Scenario	**Original Estimated Budget Cost**	**Cost Score (Original)**	**Normalized Cost Score (Original)**	**1 – (*normalized cost score - original*)**	**Estimated Budget Cost (39 eagles)**	**Normalized Cost Score (39 Eagles)**	**1 – (*normalized cost score - 39 eagles)***	**Estimated Budget Cost (79 eagles)**	**Normalized Cost Score (79 Eagles)**	**1 – (*normalized cost score - 79 eagles)***
Status Quo	$32,000	13.7	0.598	0.402	$32,000	0.138	0.862	$32,000	0.096	0.904
Comprehensive	$53,489*	22.9	1	0	$232,000**	1	0	$332,000**	1	0
No Monitoring	$0	0	0	1	$0	0	1	$0	0	1
New Metrics	$21,489*	9.2	0.402	0.598	$200,000**	0.862	0.138	$300,000**	0.904	0.096
Reduced Status Quo 1	$25,750*	11.02*	0.481	0.519	$25,750*	0.111	0.889	$25,750*	0.078	0.922
Reduced Status Quo 2	$19,500*	8.35*	0.365	0.635	$19,500*	0.084	0.916	$19,500*	0.059	0.941

*Calculated based on relative value

**Estimated based on reference values